

Duality of Caustics in Minkowski Billiards

S. Artstein-Avidan, D. Florentin, Y. Ostrover, D. Rosen

Abstract

In this paper we study convex caustics in Minkowski billiards. We show that for the Euclidean billiard dynamics in a planar smooth, centrally symmetric, strictly convex body K , for every convex caustic which K possesses, the “dual” billiard dynamics in which the table is the Euclidean unit ball and the geometry that governs the motion is induced by the body K , possesses a dual convex caustic. Such a pair of caustics are dual in a strong sense, and in particular they have the same perimeter, Lazutkin parameter (both measured with respect to the corresponding geometries), and rotation number. We show moreover that for general Minkowski billiards this phenomenon fails, and one can construct a smooth caustic in a Minkowski billiard table which possesses no dual convex caustic.

1 Introduction and Main Results

Mathematical billiard models have been studied in various contexts, including dynamical systems, ergodic theory, statistical mechanics, and geometry. Euclidean billiard dynamics concern the propagation of a point particle with no mass in some domain in \mathbb{R}^n , called the “billiard table”. The particle moves in straight lines until it encounters the boundary. There, it reflects specularly, and bounces off according to the classical reflection law: the angle of incidence equals the angle of reflection. This law is the consequence of the following variational principle: for two points a, b in the interior of the billiard table, the reflection point x on the boundary which is part of the billiard orbit between a and b , is a critical point for the Euclidean length of the broken line \overline{axb} (see e.g., [20] and [28]).

The notion of “*Minkowski billiard*” was introduced by Gutkin and Tabachnikov in [14], as an important special case of the natural extension of Euclidean billiard dynamics to the Finsler setting. In the Minkowski case, the geometry that governs the billiard dynamics is determined by a norm on \mathbb{R}^n . The reflection law for the associated Minkowski billiard follows from a variational principle analogous to the one mentioned above, where the Euclidean length is replaced by this norm. It is sometimes useful to think of the Minkowski billiard dynamics as a dynamical system associated with a pair of convex bodies $K \subset \mathbb{R}_q^n$, and $T \subset \mathbb{R}_p^n$, where the ambient space $\mathbb{R}_q^n \times \mathbb{R}_p^n$ is the classical phase space. Here, one of the bodies, say K , plays the role of the billiard table, and the other body T determines a norm, given by the so-called support function h_T , which controls the billiard dynamics in K (see Section 2.2 below for the precise formulation). The (K, T) -billiard dynamics takes place on the boundary $\partial(K \times T)$, and the projections of the associated orbits to \mathbb{R}_q^n (respectively

to \mathbb{R}_p^n) are called T -billiard trajectories in K (respectively K -billiard trajectories in T). When $T = B$ is the Euclidean unit ball, the B -billiard trajectories in K are the classical billiard orbits in K . With this description, it follows immediately that there is a one-to-one correspondence between T -billiard trajectories in K , and K -billiard trajectories in T . In what follows we call such a pair of trajectories “dual billiard trajectories” (see Section 2.2.1).

Motivated by this orbit-to-orbit duality, in this paper we study a certain caustic-to-caustic duality for planar Minkowski billiards. Recall that for Euclidean billiards, a convex *caustic* is a convex body with the property that once a billiard trajectory is tangent to it, it remains tangent after every reflection at the boundary. Caustics play an important role in the study of Euclidean billiard dynamics, and their existence has an essential impact on the dynamics (see Section 2.1 below for more details). The definition of a convex caustic in the Minkowski billiard setting, at least when the body T which induces the geometry is smooth, centrally symmetric, and strictly convex, is the same as in the Euclidean case, i.e., a T -caustic in K is a convex body in \mathbb{R}_q^n such that every T -billiard trajectory once tangent to it, stays tangent after every reflection at ∂K . Similarly, a K -caustic in T is a convex body in \mathbb{R}_p^n with the property that every T -billiard trajectory once tangent to it, remains tangent to it after every reflection with ∂T . In the non-centrally symmetric case, one must distinguish between ‘left’ and ‘right’ caustics, which for simplicity we wish to avoid in this text. Thus, we introduce the following terminology for planar Minkowski billiards.

Definition 1.1. A “symmetric billiard configuration” is a pair (K, T) of C^1 -smooth centrally symmetric strictly convex bodies, where $K \subset \mathbb{R}_q^2$, and $T \subset \mathbb{R}_p^2$.

For a symmetric billiard configuration (K, T) we denote by $\mathfrak{C}(K, T)$ the set of all the convex T -caustics in the billiard table K . The set $\mathfrak{C}(T, K)$ is defined in a similar manner.

Definition 1.2 (Caustic Duality). Let (K, T) be a symmetric billiard configuration. Two convex caustics $C \in \mathfrak{C}(K, T)$ and $C' \in \mathfrak{C}(T, K)$ are called “dual caustics” if for every T -billiard trajectory in K which is tangent to C , the dual K -billiard trajectory in T is tangent to C' , and vice versa.

Remark 1.3. We emphasize that the dual convex caustic to $C \in \mathfrak{C}(K, T)$, if exists, is unique (see Section 2.3.3). Moreover, we shall see below that several parameters naturally associated with a convex caustic $C \in \mathfrak{C}(K, T)$, such as its rotation number, perimeter, and the so-called Lazutkin parameter (where the last two quantities are defined via the corresponding support function h_T), are the same for any pair of smooth strictly convex dual caustics (see Section 2.3.4 below for the definitions and more details).

Our main result, concerning dual caustics for Euclidean billiards, is the following.

Theorem 1.4. Let $K \subset \mathbb{R}_q^2$ be a C^1 -smooth centrally symmetric strictly convex body, and $B \subset \mathbb{R}_p^2$ be the Euclidean unit disk. Then for every convex B -caustic in K there exists a dual convex K -caustic in B . Moreover, the two caustics have the same perimeter, Lazutkin parameter (both measured with respect to the corresponding geometries), and rotation number.

Remark 1.5. The dual caustic mentioned in Theorem 1.4 has a simple geometric description. More precisely, let C be a B -caustic in K . For $q \in \partial K$, denote by $e(q)$ and $b(q)$ the positive and negative tangency points to C from q , respectively (see Figure 8 below). Moreover, let $L(q) = |q - e(q)| + |q - b(q)|$. Then, the dual caustic $C' \in \mathfrak{C}(B, K)$ is given by

$$C' = \text{Conv} \left\{ \frac{e(q) - b(q)}{L(q)} : q \in \partial K \right\}.$$

In particular, if C is a polygon, then so is C' . On the other hand, if C is sufficiently smooth, then C' is smooth, and its boundary is parametrized by $q \mapsto \frac{e(q) - b(q)}{L(q)}$, where $q \in \partial K$.

The notion of caustics for planar convex billiard tables is closely related with the notion of an ‘invariant circle’ of the associated monotone twist map (see Section 2.3.3 below). In particular, any convex caustic gives rise to such an invariant circle. It follows immediately from the definition that for any symmetric billiard configuration (K, T) , there is a natural bijection between the invariant circles of the twist map associated with the T -billiard dynamics in K , and the invariant circles of the twist map associated with the K -billiard dynamics in T . However, in general, not every invariant circle corresponds to a *convex* caustic (see e.g., [13] and [19]). Thus, the existence of a convex dual caustic in Theorem 1.4 above is not obvious. Moreover, as Theorem 1.6 below demonstrates, this result fails in general if one replaces the Euclidean ball B with an arbitrary convex body.

Theorem 1.6. *There exists a symmetric billiard configuration (K, T) and a smooth strictly convex T -caustic in K which possesses no dual convex caustic, that is, there exists no K -caustic in T which is dual to it in the sense of Definition 1.2 above.*

Notations: Throughout the paper the word “smooth” means C^1 -smooth, unless explicitly stated otherwise. For a smooth function $F : \mathbb{R}^n \rightarrow \mathbb{R}$ we write ∇F for its gradient. A convex body K in \mathbb{R}^n is a compact convex set with non-empty interior. It is said to be ‘strictly convex’ if its boundary contains no line segment, and ‘centrally symmetric’ if it is symmetric with respect to the origin, i.e., $K = -K$. For a centrally symmetric convex body $K \subset \mathbb{R}^n$, the gauge function $g_K : \mathbb{R}^n \rightarrow \mathbb{R}$ given by $g_K(x) = \inf\{r \geq 0 : x \in rK\}$ defines a norm in \mathbb{R}^n , denoted by $\|\cdot\|_K$. The support function $h_K : \mathbb{R}^n \rightarrow \mathbb{R}$ of a convex body K is defined by $h_K(u) = \sup\{\langle x, u \rangle : x \in K\}$. If K is centrally symmetric, then one has $h_K(u) = \|u\|_{K^\circ}$, where $K^\circ = \{y \in \mathbb{R}^n : h_K(y) \leq 1\}$ is the dual (polar) body of K . The Euclidean unit ball is denoted by B^n (when there is no risk of confusion, we sometimes drop the superscript n to simplify notation). We denote by $d_H(\cdot, \cdot)$ the Hausdorff distance between compact sets in \mathbb{R}^n . The convex hull of the sets $\{A_i\}_{i=1}^m$ in \mathbb{R}^n is denoted by $\text{Conv}(A_1, A_2, \dots, A_m)$. An oriented line ℓ in \mathbb{R}^n is denoted by (x, v) , where $x \in \ell$ and v is the orientation of ℓ . We denote by $\bar{\ell} := (x, -v)$ the same line with the reversed orientation. Given a planar convex body $C \subset \mathbb{R}^2$, an oriented line $\ell = (x, v)$ is said to be positively tangent to C if ℓ and C intersect, and C is contained in the closed left half-plane determined by ℓ , i.e., $C \subset \{y \in \mathbb{R}^2 : \det(v, y - x) \geq 0\}$. The oriented line $\ell = (x, v)$ is said to be negatively tangent to C if the oriented line $\bar{\ell}$ is positively tangent to C . For a centrally symmetric convex body $K \subset \mathbb{R}^2$ and a (rectifiable) curve $\gamma \subset \mathbb{R}^2$, we write $\text{Length}_{h_K}(\gamma)$ for the length

of γ measured with respect to h_K (see e.g., [30]), and $\text{Length}(\gamma)$ for the Euclidean length. Finally, we denote the perimeter of a planar convex set C with respect to h_K by $\text{Per}_{h_K}(C)$, and the Euclidean perimeter by $\text{Per}(C)$.

Organization of the paper: In Section 2 we provide the necessary background on Minkowski billiards, and present two models for the dynamics: a continuous one via the characteristic foliation in the classical phase space, and a discrete-time model via a Poincaré surface of section. Then, confining ourselves to the planar case, we discuss Minkowski caustics and their relations with invariant circles of the corresponding monotone twist map, and introduce some natural invariants associated with them. In Section 3 we prove Theorem 1.4 on the existence of dual caustics for Euclidean billiard dynamics. This is done first for sufficiently regular caustics, and then, using an approximation argument, for the general case. In Section 4 we show that the situation is radically different for arbitrary Minkowski billiards, and prove Theorem 1.6. Finally, in the Appendix we sketch an alternative proof of Theorem 1.4 in the case where the caustic C is a convex polygon.

Acknowledgment: The first named author is partially supported by ISF grant number 665/15. The third named author is partially supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme, starting grant No. 637386, and by the ISF grant No. 1274/14. The fourth named author is partially supported by the European Research Council advanced grant 338809.

2 Preliminaries and Background

2.1 Some Historical Background

As mentioned in the introduction, for a Euclidean billiard table, a convex caustic[†] is a convex body with the property that every billiard trajectory once tangent to it, remains tangent after every reflection at the boundary. For example, it follows from elementary geometrical considerations that elliptical planar billiard tables have a 1-parameter family of convex caustics given by confocal ellipses. This phenomenon is closely related to the classical fact that the billiard dynamics inside an ellipse is an integrable dynamical system (see e.g., [20] and [28]). For convex billiard tables of dimension at least three, Berger [4] and Gruber [11] proved that for Euclidean billiards only ellipsoids have convex caustics. On the other hand, convex caustics do play an important role in the study of planar Euclidean billiards. For example, the presence of a convex caustic implies that the dynamics cannot be ergodic, as the corresponding invariant curve (family of rays tangent to a caustic – see Section 2.3.3 below) separates the phase space into invariant components. Caustics are also closely related with the integrability of the billiard dynamics (see e.g., [5] and [12]). Finally, in the theory of semiclassical approximations, convex caustics can be used to estimate eigenvalues, and to construct quasimodes for the corresponding Dirichlet problem (see [22]).

[†]The term “caustic” (which means burning in Greek) is borrowed from optics, where it means a curve on which light rays focus after being reflected off a mirror.

For 2-dimensional Euclidean billiards, the existence of (uncountably many) caustics near the boundary of a sufficiently smooth convex billiard table with positive curvature was first proved by Lazutkin [21] using KAM theory. Lazutkin's original proof assumes the existence of 553 continuous derivatives of the billiard table. Later, Douady [9] reduced the degree of differentiability to six. In contrast, Mather [23] showed that if the curvature of the boundary vanishes at one point, then the billiard possesses no caustics at all. A quantitative version of this result was given by Gutkin and Katok in [13]. Hubacher [16] proved that no caustics exist near a boundary whose curvature is discontinuous at some point. For further information on caustics of Euclidean planar billiard tables we refer the reader to [12, 13, 19, 22, 28, 29].

In [14], Gutkin and Tabachnikov made a first step in investigating caustics of Minkowski billiards. In particular, they extended the above mentioned result of Mather to this setting. More precisely, they showed that planar convex billiard tables whose Minkowski curvature is not strictly positive do not have caustics. To the best of our knowledge, analogous results in the Minkowski setting to the works of Lazutkin [21] on the existence of caustics near the boundary, of Gutkin-Katok [13] on caustic-free regions, and of Berger [4] and Gruber [11] regarding the non-existence of caustics in higher dimensions, are yet to be explored.

2.2 The Minkowski (K, T) -Billiard Dynamics

Minkowski billiards model, for example, the propagation of light in an anisotropic homogeneous medium, i.e., where the velocity depends on the direction, but is independent of the position (see [14], and also [17] for a motivation coming from the study of light patterns observed in liquid crystal layers). Below we present two models for Minkowski billiards. The first is a continuous description given by characteristic foliations on non-smooth convex hypersurfaces in the classical phase space $\mathbb{R}^{2n} = \mathbb{R}_q^n \times \mathbb{R}_p^n$ as described in [2]. The second model, which essentially goes back to Birkhoff [7] (and is somewhat more common in the study of billiard dynamics) is a time-discrete dynamical system which can be derived from the first one using Poincaré sections. Before we turn to the precise formulation, we give first a non-formal geometric description of the (K, T) -billiard dynamics.

As mentioned in the introduction, the Minkowski (K, T) -billiard is a dynamical system associated with a pair of convex bodies, $K \subset \mathbb{R}_q^n$ and $T \subset \mathbb{R}_p^n$. We shall assume for simplicity that K and T are C^1 -smooth convex bodies. The (K, T) -billiard dynamics takes place on the boundary $\partial(K \times T)$, and can geometrically be described as follows (see Section 2.2.1 below): suppose we start at some point $(q_0, p_0) \in K \times \partial T$. The dynamics is defined so that we move in $K \times \partial T$ from (q_0, p_0) to a point $(q_1, p_0) \in \partial K \times \partial T$, following the inner normal to the boundary ∂T at the point p_0 . When we hit the boundary ∂K at the point q_1 we start to move in $\partial K \times T$ from (q_1, p_0) to $(q_1, p_1) \in \partial K \times \partial T$ following the outer normal to ∂K at the point q_1 (see Figure 1). Next, we move from (q_1, p_1) to (q_2, p_1) following the inner normal to ∂T at p_1 , and so on and so forth. Similarly to the classical Euclidean case, T -billiard trajectories in K correspond to critical points of a length functional defined on the (infinite) cross product of the boundary ∂K , where the distances between two consecutive

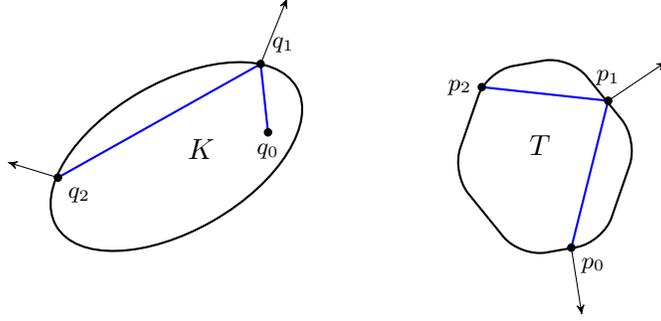


Figure 1. A (K, T) -billiard trajectory.

points are measured with respect to the support function h_T . In the case where $T = B^n$ is the Euclidean unit ball, these trajectories are the classical billiard orbits in the billiard table K . Hence, the above reflection law is a natural variation of the classical one when the Euclidean structure is replaced by a Minkowski norm. We proceed with a formal description

2.2.1 A Continuous Model

In this section we provide a continuous-time model for the Minkowski billiard dynamics, which is motivated by the natural correspondence between geodesics in a Riemannian manifold and characteristics of its unit cotangent bundle (see [2]).

Consider \mathbb{R}^{2n} with position-momentum coordinates (q, p) , and equipped with the standard symplectic form $\omega = dq \wedge dp$. Recall that the restriction of ω to a smooth closed connected hypersurface $\mathcal{S} \subset \mathbb{R}^{2n}$ defines the canonical line bundle $\mathfrak{S}_{\mathcal{S}} := \ker(\omega|_{\mathcal{S}})$, whose integral curves are called the characteristics of \mathcal{S} . Moreover, recall that the symplectic action of a closed curve γ is defined by $A(\gamma) = \int_{\gamma} \lambda$, where $\lambda = pdq$ is the Liouville 1-form.

Next, let $K \subset \mathbb{R}_q^n$, and $T \subset \mathbb{R}_p^n$ be two smooth convex bodies, and consider the g_T -unit cotangent bundle

$$U_T^*K := \{(q, p) : q \in K, \text{ and } g_T(p) \leq 1\} \simeq K \times T,$$

where g_T is the gauge function of T , and the standard inner product identifies $T^*\mathbb{R}_q^n$ with $\mathbb{R}_q^n \times \mathbb{R}_p^n$. Note that the boundary of the product $K \times T \subset \mathbb{R}_q^n \times \mathbb{R}_p^n$ is a non-smooth hypersurface. However, away from the singular strata $\partial K \times \partial T$ the canonical line bundle $\mathfrak{S}_{\partial(K \times T)}$ is well defined, and is generated by the (Hamiltonian) vector field

$$\mathfrak{X}(q, p) = \begin{cases} (-\nabla g_T(p), 0), & (q, p) \in \text{int}(K) \times \partial T, \\ (0, \nabla g_K(q)), & (q, p) \in \partial K \times \text{int}(T). \end{cases}$$

The convexity of $K \times T$ allows one to naturally extend the notion of characteristic directions to $\partial K \times \partial T$, which leads to the following definition of billiard dynamics (see [2]).

Definition 2.1. A (K, T) -billiard trajectory is the image of a piecewise smooth map* $\gamma: \mathbb{R} \rightarrow \partial(K \times T)$, $\gamma(t) = (q(t), p(t))$, such that for every $t \notin \mathcal{B}_\gamma := \{t \in \mathbb{R} : \gamma(t) \in \partial K \times \partial T\}$ one has

$$\dot{\gamma}(t) = a \mathfrak{X}(\gamma(t)), \text{ for some constant } a \in (0, \infty).$$

Moreover, for any $t \in \mathcal{B}_\gamma$, the left and right derivatives $\dot{\gamma}^\pm(t)$ of $\gamma(t)$ exist, and

$$\dot{\gamma}^\pm(t) \in \{\alpha(-\nabla g_T(p(t)), 0) + \beta(0, \nabla g_K(q(t))) : \alpha, \beta \geq 0, (\alpha, \beta) \neq (0, 0)\}.$$

Remark 2.2. For a (K, T) -billiard trajectory γ , the curve $\pi_q(\gamma)$, where $\pi_q: \mathbb{R}^{2n} \rightarrow \mathbb{R}_q^n$ is the projection onto \mathbb{R}_q^n , is called a T -billiard trajectory in K , and the curve $\pi_p(\gamma)$ (which is the corresponding projection to \mathbb{R}_p^n), is called a K -billiard trajectory in T . Thus, there is a natural one-to-one correspondence between T -billiard trajectories in K and K -billiard trajectories in T . We refer to such a pair of trajectories as “dual billiard trajectories”. Furthermore, for a pair of periodic dual trajectories, the length with respect to the support function h_T of the T -billiard trajectory in K equals the length with respect to h_K of the corresponding K -billiard trajectory in T , and also equals the symplectic action of the corresponding closed characteristic in \mathbb{R}^{2n} traversed once (see Section 7 of [14], and Section 2.4 of [2]).

Remark 2.3. The trajectories described geometrically in the beginning of this section (see Figure 1) are (K, T) -billiard trajectories in the sense of Definition 2.1. Their projections to \mathbb{R}_q^n are Minkowski T -billiard trajectories in K in the sense of [14] (which include the case of classical Euclidean billiards where $T = B$). Note also that Definition 2.1 allows also the so-called “gliding trajectories” which move solely on $\partial K \times \partial T$ (see [2] for more details). However, these trajectories will play no role in what follows. Finally, we remark that our notion of T -billiard dynamics in K includes “strange billiard orbits” in the sense of Halpern [15]. Such an orbit, when it exists, can be realized, say, as the image of a piecewise smooth map which in finite time converges to a point, and continues as a gliding trajectory. These trajectories demonstrate the following phenomenon: on the singular strata $\partial K \times \partial T$, the differential relation in Definition 2.1 might have non-unique solutions. As we shall see, this will not influence our discussion below, as away from the singular strata the differential relation has a unique solution, which, moreover, can be extended uniquely up to a time at which the solution intersects the singular strata.

2.2.2 The Discrete Model

In the previous subsection we described a continuous-time model of Minkowski billiard dynamics via the characteristic foliation in the phase space \mathbb{R}^{2n} . There exists another natural description of the billiard motion as a discrete-time dynamical system, which for Euclidean billiards goes back to Birkhoff [7] (cf. [14] for a similar approach in the Minkowski case). Consider the T -billiard dynamics in K as a mapping of pairs: $(q, v) \mapsto (q', v')$, where $q \in \partial K$, the vector v is the direction of propagation, q' is the following impact point, that is q' is the point where the trajectory starting at q with velocity v hits the boundary ∂K

*Here by piecewise smooth we mean a continuous map which is C^1 except at a discrete subset of \mathbb{R} .

next, and v' is the reflected velocity, i.e., the direction of propagation after the reflection at the point q' . More precisely, let $K \subset \mathbb{R}_q^n$ and $T \subset \mathbb{R}_p^n$ be two smooth centrally symmetric strictly convex bodies. Let $n_K(q)$ be the unit outer normal to K at the point $q \in \partial K$, and similarly for $n_T(p)$. Note that $n_K(q) = \nabla g_K(q)/|\nabla g_K(q)|$. Denote

$$\begin{aligned} (\partial K \times \partial T)_+ &= \{(q, p) \in \partial K \times \partial T : \langle n_K(q), n_T(p) \rangle > 0\}, \\ (\partial K \times \partial T)_- &= \{(q, p) \in \partial K \times \partial T : \langle n_K(q), n_T(p) \rangle < 0\}, \\ (\partial K \times \partial T)_0 &= \{(q, p) \in \partial K \times \partial T : \langle n_K(q), n_T(p) \rangle = 0\}, \end{aligned}$$

and consider the decomposition

$$\partial K \times \partial T = (\partial K \times \partial T)_+ \cup (\partial K \times \partial T)_0 \cup (\partial K \times \partial T)_-.$$

It follows from Definition 2.1 (see also [2]) that a (K, T) -billiard trajectory $(q(t), p(t))$ for which $(q(t_0), p(t_0)) \in (\partial K \times \partial T)_-$ will continue to a point in $\partial K \times \text{int}(T)$, i.e., one has $(q(t_0 + \varepsilon), p(t_0 + \varepsilon)) \in \partial K \times \text{int}(T)$, for small enough $\varepsilon > 0$, and was in $\text{int}(K) \times \partial T$ before the collision with the singular set $\partial K \times \partial T$, i.e., $(q(t_0 - \varepsilon), p(t_0 - \varepsilon)) \in \text{int}(K) \times \partial T$, for small enough $\varepsilon > 0$. In a similar manner, a (K, T) -billiard trajectory $(q(t), p(t))$ for which $(q(t_0), p(t_0)) \in (\partial K \times \partial T)_+$ arrives to this point from $\partial K \times \text{int}(T)$, and continues to move in $\text{int}(K) \times \partial T$. On the other hand, trajectories for which $(q(t_0), p(t_0)) \in (\partial K \times \partial T)_0$ may exhibit a more complicated behavior. They can move in the singular strata $\partial K \times \partial T$ as gliding trajectories, or, as the example of “strange (Euclidean) billiards” (see [15]) shows, the point $(q(t_0), p(t_0))$ might be the accumulation point of the collision points of a billiard trajectory which has an infinite number of collisions in finite time. We remark that under sufficient smoothness conditions on the boundaries ∂K and ∂T , strange billiards cannot occur, as shown in Proposition 2.12 in [2].

Next, consider the Poincaré return map associated with the subset

$$\Sigma := (\partial K \times \partial T)_+ \cup (\partial K \times \partial T)_-,$$

which the (K, T) -billiard trajectories cross transversally. By the above discussion, the return map is indeed well defined. We shall denote this map by $\Psi = \Psi_{K, T}$. One can check directly that Ψ interchanges the two sets $(\partial K \times \partial T)_+$ and $(\partial K \times \partial T)_-$, i.e.,

$$\Psi : (\partial K \times \partial T)_+ \rightarrow (\partial K \times \partial T)_- \text{ and } \Psi : (\partial K \times \partial T)_- \rightarrow (\partial K \times \partial T)_+.$$

We observe that Σ is a symplectic submanifold of \mathbb{R}^{2n} , and since the vector field \mathfrak{X} generates the null direction of ω on Σ , the Poincaré map Ψ is symplectic, i.e.,

$$\Psi^*(dq \wedge dp) = dq \wedge dp. \tag{1}$$

Considering the dynamics of the (K, T) -billiard trajectories given in Definition 2.1 above, one sees that for $(q, p) \in (\partial K \times \partial T)_+$ one has $\Psi(q, p) = (q', p)$, where q' is the next impact point of the ray in K emanating from q in direction $-n_T(p)$ with the boundary ∂K . Note that this point exists, as $n_T(p)$ points outside of K at q by the definition of the set $(\partial K \times \partial T)_+$. Similarly, for $(q, p) \in (\partial K \times \partial T)_-$, one has $\Psi(q, p) = (q, p')$, where p' is the next impact point of the ray in T emanating from p in the direction $n_K(q)$.

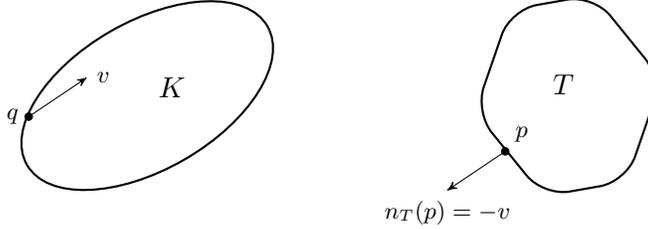


Figure 2. The isomorphism $\mathcal{P}(K) \simeq (\partial K \times \partial T)_+$, $(q, v) \mapsto (q, p)$.

Definition 2.4. *In the following we refer to Ψ^2 as the “discrete (K, T) -billiard map”.*

To motivate this, consider the restriction of Ψ^2 to, say, $(\partial K \times \partial T)_+$. Note that the set $(\partial K \times \partial T)_+$ can be viewed as the phase space of the T -billiard dynamics in K . Indeed, a natural description of the phase space associated with Minkowski billiards (see [14]) is the set of inward pointing h_T -unit tangent vectors with foot at ∂K , i.e.,

$$\mathcal{P}(K) := \{(q, v) : q \in \partial K, v \in T_q K, h_T(v) = 1, \langle v, n_K(q) \rangle < 0\}.$$

There is a natural bijection $\mathcal{P}(K) \rightarrow (\partial K \times \partial T)_+$ via $(q, v) \mapsto (q, p)$, where $v = -n_T(p)$ (see Figure 2). Under this identification, the map Ψ^2 takes the form of the usual billiard ball map $(q, v) \mapsto (q', v')$, as described above. A similar identification as above holds for the set $(\partial K \times \partial T)_-$ as the phase space $\mathcal{P}(T)$ of the K -billiard dynamics in T .

2.3 Planar Minkowski Billiards

In this section we restrict attention to the case of convex bodies in the plane. In this case, a convex set determines in a simple geometric way a 1-parameter family of convex billiard tables for which it is a Minkowski caustic. This geometric construction will play a central role in what follows, and will be described in Section 2.3.1. In Section 2.3.2 we recall the fact that the Minkowski billiard map is a monotone twist map, and in Section 2.3.3 we discuss the relation between invariant circles of this monotone twist map and convex caustics. In particular, we rephrase the notion of caustic duality via duality of the corresponding invariant circles. Finally, in Section 2.3.4 we discuss some classical invariants of caustics and their behaviour under caustic duality.

2.3.1 The String Construction

It is well known that for Euclidean billiards, for every convex set C in the plane one can associate a 1-parametric family of billiard tables $(K_L)_{L > \text{Per}(C)}$ such that each table has C as a caustic. Roughly speaking, ∂K_L is obtained by the following procedure: wrap a loop of inelastic string of length L around C . Then, pull the string tight away from C to produce a point p on the boundary of the billiard table K_L . Finally, move the point p around C , keeping the string tight, to obtain the rest of ∂K_L . This technique is known as

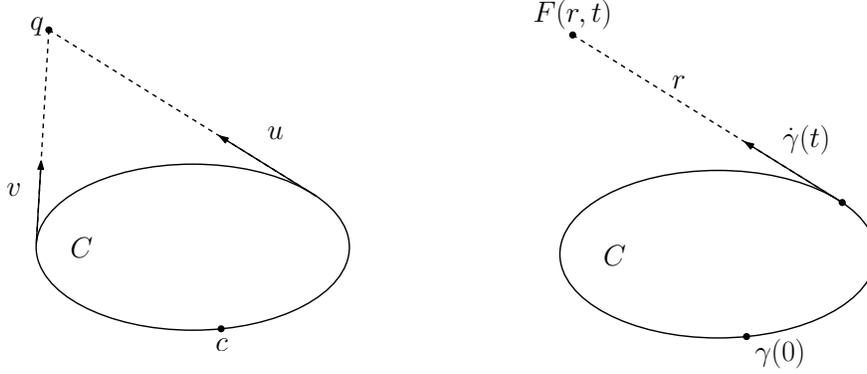


Figure 3

the “gardener’s string construction” (see e.g., [8, 29]). Note that when this construction is applied to a closed line segment one obtains an ellipse. For an illustration of a Euclidean string construction over a triangle see Figure 10 in the Appendix.

This string construction can be naturally generalized to planar Minkowski billiards (see Section 3 of [14]), where lengths are measured with respect to some Minkowski norm on \mathbb{R}^2 . More precisely, let $C \subset \mathbb{R}_q^2$ be a convex set, and $T \subset \mathbb{R}_p^2$ a centrally symmetric convex body. The result of the h_T -string construction over C with string length $L > \text{Per}_{h_T}(C)$ is, by definition,

$$K = \{q \in \mathbb{R}^2 : \text{Per}_{h_T}(\text{Conv}(q, C)) \leq L\}.$$

Remark 2.5. The quantity $L - \text{Per}_{h_T}(C)$ is called the “Lazutkin parameter” associated with the above string construction over C (see e.g., Chapter 3 in [27] or [14]).

Note that the boundary ∂K is a level set of the function $f(q) = \text{Per}_{h_T}(\text{Conv}(q, C))$. The following lemma (cf. the proof of Lemma 3.6 in [14]) shows that the function f is smooth. For the proof we shall need some basic facts from the differential geometry of curves in normed planes. We refer the reader to [3, 6, 24] for the relevant definitions and properties.

Lemma 2.6. *Let $C \subset \mathbb{R}_q^2$ be a compact convex set, and $T \subset \mathbb{R}_p^2$ a centrally symmetric strictly convex body. Then, the function $f(q) = \text{Per}_{h_T}(\text{Conv}(q, C))$ is C^1 -smooth on $\mathbb{R}_q^2 \setminus C$. Moreover, one has*

$$\nabla f(q) = n_{T^\circ}(u) + n_{T^\circ}(v),$$

where u and v are h_T -unit vectors parallel to the tangent lines from q to C , pointing towards the point q (see Figure 3).

Proof of Lemma 2.6. We first assume that T° is C^2 -smooth, and that the convex set C is C^2 -smooth, with non-vanishing Minkowski curvature with respect to the norm h_T (see [3, 6, 24]). We remark that the condition of “non-vanishing Minkowski curvature” is independent of the choice of the Minkowski norm. Let $q \in \mathbb{R}^2 \setminus C$. We pick a point $c \in \partial C$ on the “other side” of ∂C . We can write, near q , $f(q) = l_1(q) + l_2(q)$, where $l_1(q)$ and $l_2(q)$

are the lengths of the shortest paths from c to q in $\mathbb{R}^2 \setminus C$, in the counter-clockwise and clockwise directions, respectively (see Figure 3). We will compute the derivative of l_1 . To this end, let $\gamma: [0, A] \rightarrow \partial C$ be a counter-clockwise, h_T -unit parameterization of ∂C such that $\gamma(0) = c$. We consider the local diffeomorphism

$$F(r, t) = \gamma(t) + r\gamma'(t), \quad t \in [0, A], r > 0.$$

In these coordinates, l_1 is particularly simple, and one has that $F^*l_1(r, t) := l_1 \circ F(r, t) = r + t$. Evidently, F^*l_1 is C^1 -smooth, and $F^*dl_1 = dF^*l_1 = dr + dt$. Writing this out in Cartesian coordinates, we get

$$DF^T(r, t) \cdot \nabla l_1(F(r, t)) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (2)$$

To differentiate F , we recall the Frenet equations in a normed plane (see e.g. [3]). By definition, $\gamma'(t)$ lies on the boundary of the h_T -unit ball, which as noted above, is the polar body T° . The right normal vector $n_\gamma(t)$ to γ at t is by definition the counter-clockwise tangent to T° at the point $\gamma'(t)$, unit normalized in the so-called anti-norm h_{JT} , where J is a counter-clockwise rotation of the plane by 90° . Equivalently, $-Jn_\gamma(t) = n_{T^\circ}(\gamma'(t))$. The (first) Frenet equation for γ reads $\gamma''(t) = k_m(t)n_\gamma(t)$, where $k_m(t)$ is the Minkowski curvature. Using this we compute

$$DF(r, t) = \left(\gamma'(t) + rk_m(t)n_\gamma(t), \gamma'(t) \right).$$

Inverting the matrix $DF(r, t)^T$ and using (2), we have

$$\nabla l_1(F(r, t)) = -Jn_\gamma(t) = n_{T^\circ}(\gamma'(t)).$$

Observe that in the notations of the lemma, $\gamma'(t) = u$, thus we have computed $\nabla l_1(q) = n_{T^\circ}(u)$. An identical computation gives $\nabla l_2(q) = n_{T^\circ}(v)$ which, since $f(q) = l_1(q) + l_2(q)$, proves the claim in this case.

Now we turn to the case of a general compact convex set C . Let C_j be a sequence of C^2 -smooth convex bodies with positive Minkowski curvature which converge to C in the Hausdorff topology. Consider the associated string length functions $f_j(q) = \text{Per}_{h_T}(\text{Conv}(q, C_j))$. One easily verifies that the sequence (f_j) tends to f uniformly outside a neighbourhood of C , and by the previous step,

$$\nabla f_j(q) = n_{T^\circ}(u_j) + n_{T^\circ}(v_j),$$

where u_j and v_j are h_T -unit vectors parallel to the two tangents from q to C_j . Clearly $u_j \rightarrow u$ and $v_j \rightarrow v$ point-wise. The function f is convex, and hence has a well-defined sub-differential $\partial f(q)$ at every point (see e.g. Chapter 1 in [26]). Moreover, at a point where the sub-differential is a singleton the function f is differentiable. Since $f_j \rightarrow f$, one has that for every q ,

$$n_{T^\circ}(u) + n_{T^\circ}(v) = \lim_{j \rightarrow \infty} \nabla f_j(q) \in \partial f(q).$$

Consider a level set $S := \{q : f(q) = L\}$. At each $q \in S$, the vector $n(q) := n_{T^\circ}(u) + n_{T^\circ}(v)$ belongs to the sub-differential $\partial f(q)$, where the latter is a base of the normal cone for the level set S (see again [26]). The function $n(q)$ is a continuous section of this normal cone. However, normal cones at distinct points of S have disjoint relative interiors (Chapter 2 in [26]). This clearly implies that the sub-differentials are all singletons, and so f is differentiable. Moreover, it follows that $\nabla f(q) = n(q)$, whence f is C^1 , as claimed. This completes the proof when T° is C^2 -smooth. To remove the requirement on the existence of the second derivative, we approximate T° by a sequence of C^2 -smooth convex bodies, and observe that all the needed properties are preserved for the above argument. \square

Lemma 2.7. *Let K be the body formed by an h_T -string construction over a convex body C . Then K is convex. Moreover, if T is smooth and strictly convex, then so is K .*

In the proof (and in the sequel) we shall use the following well-known fact from convex geometry: for a centrally symmetric convex body $T \subset \mathbb{R}^2$, the h_T -perimeter of a planar convex body D is the mixed area of the body and T rotated by 90° (see, e.g., Section 2.3 in [30]). More precisely, denote by $V(\cdot, \cdot)$ the mixed area function, and let JT be the body obtained by rotation of T by 90° . Then,

$$\text{Per}_{h_T}(D) = 2V(D, JT). \quad (3)$$

Proof of Lemma 2.7. Mixed areas are known to be linear (with respect to Minkowski sum with positive coefficients) and monotone in each argument (see [30] Section 2.3). Hence, relation (3) implies that $\text{Per}_{h_T}(D)$ is linear and monotone with respect to D . Thus, for $x, y \in K$, since for $\lambda \in (0, 1)$

$$\text{Conv}((1 - \lambda)x + \lambda y, C) \subseteq (1 - \lambda)\text{Conv}(x, C) + \lambda\text{Conv}(y, C), \quad (4)$$

one has that

$$\begin{aligned} \text{Per}_{h_T}(\text{Conv}((1 - \lambda)x + \lambda y, C)) &\leq \text{Per}_{h_T}((1 - \lambda)\text{Conv}(x, C) + \lambda\text{Conv}(y, C)) \\ &= (1 - \lambda)\text{Per}_{h_T}(\text{Conv}(x, C)) + \lambda\text{Per}_{h_T}(\text{Conv}(y, C)) \quad (5) \\ &\leq L. \end{aligned}$$

Thus, $(1 - \lambda)x + \lambda y$ is also in K , which proves the convexity of K . Moreover, note that the inclusion in (4) is in fact strict when $x \neq y$. Indeed, the point $(1 - \lambda)x + \lambda z$ is not in the left-hand side of (4), for one of the tangency points z from $(1 - \lambda)x + \lambda y$ to C . If T is smooth, the surface area measure dS_{JT} (see Definition 2.3.12 in [30]) of JT has full support, and from the fact that $2V(D, JT) = \int_{S^1} h_D dS_{JT}$ (see *ibid.*), it follows that the first inequality in (5) is strict, and hence K is strictly convex. Finally, if T is strictly convex, then by Lemma 2.6, the boundary ∂K is a regular level set of the C^1 -function $f(q) = \text{Per}_{h_T}(\text{Conv}(q, C))$, which completes the proof of the lemma. \square

Another important consequence of Lemma 2.6, which appeared in [14], is that those billiard tables admitting a convex T -caustic C are precisely all possible h_T -string constructions over C .

Lemma 2.8 (Lemma 3.6 in [14]). *Let (K, T) be a symmetric billiard configuration. A convex set $C \subset \mathbb{R}_q^2$ is a T -caustic in K if and only if the function $L_{h_T}(x) = \text{Per}_{h_T}(\text{Conv}(x, C))$, defined for $x \in \partial K$, is constant.*

In the sequel we shall also need the following simple facts regarding the Minkowski string construction.

Lemma 2.9. *Let C be a convex body in \mathbb{R}_q^2 , and denote by $K_{T,L}(C) \subset \mathbb{R}_q^2$ the billiard table obtained by means of a h_T -string construction with string length $L > \text{Per}_{h_T}(C)$.*

(i) *If $C \subseteq C'$ are two convex bodies, then $K_{T,L}(C') \subseteq K_{T,L}(C)$,*

(ii) *If $L_1 \leq L_2$, then $K_{T,L_1}(C) \subseteq K_{T,L_2}(C)$,*

(iii) *If $T_1 \subseteq T_2$ are two convex bodies, then $K_{T_1,L}(C) \subseteq K_{T_2,L}(C)$.*

Proof of Lemma 2.9. (i) Take a point $x \in K_{T,L}(C')$, so that the $\text{Per}_{h_T}(\text{Conv}(x, C')) \leq L$. Clearly, $\text{Conv}(x, C) \subseteq \text{Conv}(x, C')$, and so by the monotonicity of the h_T -perimeter (which follows, say, from (3) above), one has that $\text{Per}_{h_T}(\text{Conv}(x, C)) \leq L$. Thus, $x \in K_{T,L}(C)$. Items (ii) and (iii) are immediate from the definition. \square

2.3.2 The Monotone Twist Condition

In this section we verify that the Minkowski billiard map is a monotone twist map. This description will be useful later for relating convex caustics and invariant circles, and when we discuss various classical parameters associated with convex caustics (see Section 2.3.3 below). For Euclidean billiard dynamics this is a classical fact (see e.g., [27] and [28]), and for Minkowski billiards the details can be found in [14]. Below we use somewhat different terminology and notations from [14], and therefore we repeat the argument for the sake of completeness.

Let (K, T) be a symmetric billiard configuration. We introduce the following notations: denote $S^1 = S_P^1 = \mathbb{R}/P\mathbb{Z}$, where $P = \text{Per}_{h_T}(K)$. Note that according to formula (3) one has $P = \text{Per}_{h_T}(K) = \text{Per}_{h_K}(T)$. Let $\gamma_K: S^1 \rightarrow \partial K$ be a h_T -unit speed, counter-clockwise parametrization of ∂K . For $q \in \partial K$ let $\tau_K(q)$ denote the positive h_T -unit tangent to ∂K , i.e., $\dot{\gamma}(t) = \tau_K(\gamma(t))$. This gives the following natural identification of annuli: set $A_K := S^1 \times (-1, 1)$, then one has

$$(\partial K \times \partial T)_+ \simeq A_K, \quad (q, p) \mapsto (t, s), \quad \text{where } q = \gamma_K(t) \text{ and } s = \langle -p, \tau_K(q) \rangle. \quad (6)$$

In a similar way $(\partial K \times \partial T)_-$ is identified with $A_T := S^1 \times (-1, 1)$ via the map

$$(q, p) \mapsto (t, s), \quad \text{where } p = \gamma_T(t) \text{ and } s = \langle q, \tau_T(p) \rangle.$$

Let us first verify that the map in (6) is a well-defined bijection. For $q \in \partial K$, denote

$$\partial T_{+,q} = \{p \in \partial T : \langle n_K(q), n_T(p) \rangle > 0\}.$$

Lemma 2.10. *With the above notations, for every $p \in \partial T_{+,q}$ one has $\langle p, \tau_K(q) \rangle \in (-1, 1)$. Moreover, the map*

$$\partial T_{+,q} \rightarrow (-1, 1), \quad \text{given by } p \mapsto s = \langle -p, \tau_K(q) \rangle,$$

is a (monotone) bijection.

Proof of Lemma 2.10. First, as $\tau_K(q) \in \partial T^\circ$ (note that by definition $h_T(\tau_K(q)) = 1$), one has $|\langle p, \tau_K(q) \rangle| \leq 1$. Moreover, for every $p \in \partial T$ and $v \in \partial T^\circ$ one has $\langle p, v \rangle = \pm 1$ if and only if $v = \pm n_T(p)$. In particular, if $\langle p, \tau_K(q) \rangle = \pm 1$, then $\langle n_K(q), n_T(p) \rangle = \langle n_K(q), \pm \tau_K(q) \rangle = 0$, which proves the first assertion. For the second, observe that the map

$$\partial T_{+,q} \rightarrow (-1, 1), \quad \text{given by } p \mapsto \langle p, \tau_K(q) \rangle,$$

is continuous and monotone, as its derivative along ∂T is $\langle \tau_T(p), \tau_K(q) \rangle$, which is positive by the definition of $\partial T_{+,q}$. \square

Let us note that under the identification $(\partial K \times \partial T)_+ \simeq A_K$ given by (6) the 1-form sdt on A_K pulls back to the Liouville form $-pdq$ on $(\partial K \times \partial T)_+$, and so $ds \wedge dt$ pulls back to the standard symplectic form $dq \wedge dp$. One can naturally pull back the discrete (K, T) -billiard map Ψ^2 (see Definition 2.4) restricted to $(\partial K \times \partial T)_+$ to a bijective map ϕ_K from A_K to itself. Moreover, we lift ϕ_K to a map $\tilde{\phi}_K$ on $\mathbb{R} \times (-1, 1)$ satisfying $\tilde{\phi}_K(r + P, s) = \tilde{\phi}_K(r, s) + (P, 0)$. Note that one can also continuously extend $\tilde{\phi}_K$ to $\mathbb{R} \times [-1, 1]$ so that on the boundaries one has $\tilde{\phi}_K(r, 1) = (r + P, 1)$ and $\tilde{\phi}_K(r, -1) = (r, -1)$. The map $\tilde{\phi}_K$ satisfies an important condition: it is a ‘‘monotone twist map’’ (see e.g., Definition 1.1.1 in [27]), with generating function h given by

$$h(r, r') = -h_T(\gamma_K(r) - \gamma_K(r')), \quad (7)$$

where γ_K has been lifted to a P -periodic function on \mathbb{R} . To show this, one computes

$$\frac{\partial h}{\partial r} = \langle -n_K(q - q'), \tau_K(q) \rangle,$$

where $q = \gamma_K(r)$ and $q' = \gamma_K(r')$, and similarly

$$\frac{\partial h}{\partial r'} = \langle n_K(q - q'), \tau_K(q) \rangle.$$

Using the description of the billiard dynamics in Section 2.2.2, one easily sees that

$$\tilde{\phi}_K(r, s) = (r', s') \iff \begin{cases} \frac{\partial h}{\partial r} = s, \\ \frac{\partial h}{\partial r'} = -s'. \end{cases}$$

In particular, $\tilde{\phi}_K$ preserves the area form $dr \wedge ds$. The monotone twist condition for this mapping follows from Lemma 2.10. Indeed, the twist condition is equivalent to the fact that for a given q the set $\partial T_{+,q}$ is mapped onto the interval $(-1, 1)$ in a monotone manner.

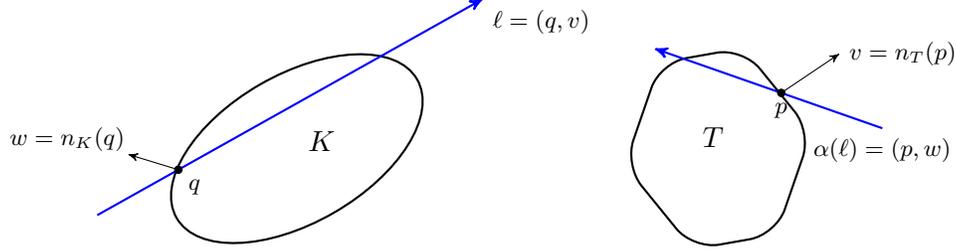


Figure 4. The (K, T) -duality transform $(q, v) \xrightarrow{\alpha} (p, w)$.

2.3.3 Minkowski Caustics and Invariant Circles

We recall from the previous section that the discrete T -billiard map $\phi_K : A_K \rightarrow A_K$ associated with the table $K \subset \mathbb{R}_q^2$ is a monotone twist map. An invariant circle of ϕ_K is an embedded circle in A_K which is homotopically non-trivial and ϕ_K -invariant. A classical result by Birkhoff states that any invariant circle of a monotone twist map is the graph of a Lipschitz function from S_P^1 to $(-1, 1)$ (see Section 1.3 in [27] and the references therein). We note that under the identification $(\partial K \times \partial T)_+ \simeq A_K$ (6), a ϕ_K -invariant circle corresponds to an embedded circle in $(\partial K \times \partial T)_+ \subset \partial K \times \partial T$, which is homotopic to either one of the two disjoint circles forming $(\partial K \times \partial T)_0$, and invariant under Ψ^2 .

Any convex T -caustic C in K gives rise to a ϕ_K -invariant circle, as follows. For $q \in \partial K$, consider the positive tangent line to C from q with direction v . The set of all such pairs $(q, v) \in \mathcal{P}(K) \simeq (\partial K \times \partial T)_+$ forms an embedded circle $\Gamma \subset A_K$, which by construction is a graph over S^1 and, in particular, not contractible. Moreover, as C is a caustic, Γ is ϕ_K -invariant. We emphasize that the converse is false in general, i.e., not every invariant circle corresponds to a *convex* caustic (see e.g., [13, 19]). Note that given a ϕ_K -invariant circle $\Gamma \subset A_K \simeq (\partial K \times \partial T)_+$, its image under Ψ defines a ϕ_T -invariant circle $\Psi(\Gamma) \subset A_T \simeq (\partial K \times \partial T)_-$. This fact allows us to rephrase Definition 1.2 in terms of invariant circles.

Corollary 2.11. *Let (K, T) be a symmetric billiard configuration in $\mathbb{R}_q^2 \times \mathbb{R}_p^2$. Consider two convex caustics $C \in \mathfrak{C}(K, T)$ and $C' \in \mathfrak{C}(T, K)$, with corresponding invariant circles $\Gamma \subset A_K$ and $\Gamma' \subset A_T$. Then, C and C' are dual caustics if and only if $\Psi(\Gamma) = \Gamma'$.*

Next we describe a geometric point of view on caustic duality. Let (K, T) as above. We consider the space

$$\mathcal{L}(K) = \{\ell \subset \mathbb{R}^2 \text{ an oriented line} : \ell \cap \text{int}(K) \neq \emptyset\}.$$

Observe that $\mathcal{L}(K)$ is naturally identified with the phase space, $\mathcal{P}(K) \simeq (\partial K \times \partial T)_+$, of the T -billiard dynamics in K , where the oriented line $\ell \in \mathcal{L}(K)$ is identified with the pair $(q, v) \in \mathcal{P}(K)$ such that v is the h_T -unit vector in the direction of ℓ , and $q \in \partial K \cap \ell$ is the first intersection point, that is, $q + \epsilon v \in K$ for small enough $\epsilon > 0$. One has a similar identification $\mathcal{L}(T) \simeq \mathcal{P}(T)$. Note that the identification $\mathcal{L}(K) \simeq \mathcal{P}(K)$ induces a natural topology on $\mathcal{L}(K)$. From these identifications, the Poincaré map induces a map, which (by

a slight abuse of notation) we also denote by $\Psi: \mathcal{L}(K) \rightarrow \mathcal{L}(T)$. Its square, the billiard ball map Ψ^2 , induces a map $\mathcal{L}(K) \rightarrow \mathcal{L}(K)$, given by $\ell \mapsto \ell'$, such that the oriented segments $\ell \cap K$ and $\ell' \cap K$ form consecutive segments of a T -billiard trajectory in K .

Using the identifications $\mathcal{L}(K) \simeq \mathcal{P}(K)$ and $\mathcal{L}(T) \simeq \mathcal{P}(T)$ implicitly, we define:

Definition 2.12. *The (K, T) -duality map $\alpha = \alpha_{K, T}: \mathcal{L}(K) \rightarrow \mathcal{L}(T)$ is given by*

$$\alpha(q, v) = (p, w), \text{ where } v = n_T(p) \text{ and } w = n_K(q) \text{ (see Figure 4).}$$

The duality map $\alpha = \alpha_{T, K}: \mathcal{L}(T) \rightarrow \mathcal{L}(K)$ is defined in a similar manner. The composition of these two maps, which (by a slight abuse of notation) we denote by α^2 , is the identity map both on $\mathcal{L}(K)$ and on $\mathcal{L}(T)$. The next result shows that the maps Ψ and α are essentially the same.

For an oriented line ℓ we denote by $-\ell$ the line obtained by reflecting ℓ about the origin, and by $\bar{\ell}$ the line ℓ with the reversed orientation. We note that these two operations commute, so the notation $-\bar{\ell}$ is unambiguous.

Lemma 2.13. *Let (K, T) be a symmetric billiard configuration. Then for $\ell \in \mathcal{L}(K)$ one has $\alpha(\ell) = \Psi(\bar{\ell})$, and for $\ell \in \mathcal{L}(T)$ one has $\alpha(\ell) = \Psi(-\bar{\ell})$.*

Proof of Lemma 2.13. We shall prove the first assertion, the proof of the second being similar. Let $\ell = (q, v) \in \mathcal{L}(K)$. Consider the composition $\alpha \circ \Psi: \mathcal{L}(K) \rightarrow \mathcal{L}(K)$. By definition, $\Psi(\ell) = (p, w) \in \mathcal{L}(T)$, for p, w satisfying $v = -n_T(p)$ and $w = n_K(q')$, where q' is the next impact point of the oriented line ℓ with ∂K . Then $\alpha(\Psi(\ell)) = (q_1, v_1)$, where q_1 and v_1 satisfy $v_1 = n_T(p)$ and $w = n_K(q_1)$. Hence, $v_1 = -v$ and $q_1 = q'$. That is, $\alpha \circ \Psi(\ell) = (q', -v) = \bar{\ell}$. As α^2 is the identity, one has $\Psi(\ell) = \alpha(\bar{\ell})$. \square

An immediate corollary of Lemma 2.13 which will be used later is the following.

Corollary 2.14. *Suppose that (K_n, T_n) is a sequence of symmetric billiard configurations which converges in the Hausdorff topology to a symmetric billiard configuration (K, T) . Suppose moreover that $\ell_n \in \mathcal{L}(K_n)$ is a sequence of oriented lines which converges to an oriented line $\ell \in \mathcal{L}(K)$. Then $\Psi_{K_n, T_n}(\ell_n) \rightarrow \Psi_{K, T}(\ell)$.*

Proof of Corollary 2.14. This follows immediately from Proposition 2.13 and the analogous continuity property of the duality map α , which holds by the fact that if D_n is a sequence of convex bodies converging (in the Hausdorff topology) to D , and $x_n \in \partial D_n$ converge to $x \in \partial D$, then $n_{D_n}(x_n) \rightarrow n_D(x)$. \square

Using the above identifications $\mathcal{L}(K) \simeq A_K$ and $\mathcal{L}(T) \simeq A_T$, one may consider the (K, T) -duality transform α as a map $A_K \rightarrow A_T$. Thus, another immediate corollary of Lemma 2.13 is the following reformulation of Corollary 2.11:

Corollary 2.15. *Let $(K, T) \subset \mathbb{R}_q^2 \times \mathbb{R}_p^2$ be a symmetric billiard configuration. The (K, T) -duality transform α maps invariant circles to invariant circles. Moreover, given two convex caustics $C \in \mathfrak{C}(K, T)$ and $C' \in \mathfrak{C}(T, K)$ with corresponding invariant circles $\Gamma \subset A_K$ and $\Gamma' \subset A_T$, the caustics C and C' are dual if and only $\alpha(\Gamma) = \bar{\Gamma}' := \{\bar{\ell}' : \ell' \in \Gamma'\}$.*

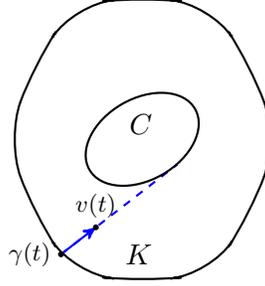


Figure 5

2.3.4 Parameters of Caustics

Here we consider some natural parameters associated with caustics, namely, the Lazutkin parameter, the perimeter, and the rotation number. We show that under mild assumptions these parameters coincide for two dual caustics. In fact, these are all marked length spectrum invariants (see e.g., Section 3.2 in [27]), and from the sequel it follows that the monotone twist maps ϕ_K and ϕ_T , for a symmetric billiard configuration (K, T) , have identical marked length spectrum.

We start by recalling the notion of rotation number associated with an invariant circle Γ of a monotone twist map $\phi: A \rightarrow A$ (see e.g., Chapter 13 in [18], or Chapter 1 in [27]). The restriction $\phi|_{\Gamma}: \Gamma \rightarrow \Gamma$ defines an orientation preserving homeomorphism of S^1 , which has a well-defined Poincaré rotation number $\omega \in [0, 1)$. The rotation number of Γ is by definition ω . It can be computed as

$$\omega = \lim_{n \rightarrow \infty} \frac{r_n}{n},$$

where $(r_n, s_n) = \tilde{\phi}^n(r_0, s_0) \in \mathbb{R} \times (-1, 1)$ is the trajectory of an arbitrary point (r_0, s_0) lying on the lift of Γ . The rotation number of a convex caustic is defined to be the rotation number of the associated invariant circle. We recall from Section 2.3.1 that to any convex T -caustic C in K one can associate its Lazutkin parameter (see Remark 2.5). Finally, by the perimeter of the T -caustic C we mean the h_T -perimeter of C .

Proposition 2.16. *Suppose that the symmetric billiard configuration (K, T) admits a pair of dual convex caustics $C \in \mathfrak{C}(K, T)$ and $C' \in \mathfrak{C}(T, K)$. Then the caustics C and C' have the same rotation number. Moreover, if we assume further that C and C' are C^1 -smooth and strictly convex, then C and C' have equal Lazutkin parameters, and perimeters.*

To prove Proposition 2.16 we find formulas to compute these parameters. We begin with the following proposition (see Theorem 3.2.10 in [27] for the Euclidean case, cf. [1]).

Proposition 2.17. *Let $C \subset K$ be a C^1 -smooth, strictly convex T -caustic, and let $\Gamma \subset A_K$ be the corresponding invariant circle. Then the h_T -perimeter of C is given by*

$$\text{Per}_{h_T}(C) = - \int_{\Gamma} s dt,$$

where s is the function appearing in the graph representation $\Gamma = \{(t, s(t)) : t \in S^1\}$.

Proof of Proposition 2.17. Using the identification $A_K \simeq (\partial K \times \partial T)_+$ one has that $\Gamma = \{(q, p(q)) : q \in \partial K\}$. Then the oriented line corresponding to Γ emanating from $q \in \partial K$ has direction $v(q) = n_T(p(q))$. Next, let $\gamma(t)$ be an h_T -unit speed parametrization of ∂K , i.e., $\dot{\gamma}(t) = \tau_K(\gamma(t))$ (where $h_T(\tau_K(q)) = 1$), and abbreviate $v(t) = v(\gamma(t))$. By the definition of a convex caustic, the line $\gamma(t) + \mathbb{R}_+v(t)$ is tangent to C . Therefore, for some $\lambda(t) > 0$, the point $\gamma(t) + \lambda(t)v(t)$ belongs to the boundary of C . We thus get a parametrization of ∂C in the form $C(t) = \gamma(t) + \lambda(t)v(t)$ (see Figure 5). Hence,

$$\dot{C}(t) = \tau_K(\gamma(t)) + \dot{\lambda}(t)v(t) + \dot{v}(t)\lambda(t). \quad (8)$$

Note that as $\dot{C}(t) \parallel v(t)$ (since the line $\gamma(t) + \mathbb{R}_+v(t)$ is tangent to the caustic C),

$$h_T(\dot{C}(t)) = \langle \dot{C}(t), n_{T^\circ}(v(t)) \rangle. \quad (9)$$

Combining relations (8) and (9), and since $n_{T^\circ}(v(t)) = p(t)$, one obtains

$$h_T(\dot{C}(t)) = \langle \tau_K(\gamma(t)), p(t) \rangle + \dot{\lambda}(t)\langle v(t), n_{T^\circ}(v(t)) \rangle + \lambda(t)\langle \dot{v}(t), n_{T^\circ}(v(t)) \rangle.$$

Note that $v(t) \in \partial T^\circ$, and thus $\langle v(t), n_{T^\circ}(v(t)) \rangle = h_T(v(t)) = 1$. Moreover, since $\dot{v}(t)$ is tangent to ∂T° , one has $\langle \dot{v}(t), n_{T^\circ}(v(t)) \rangle = 0$. Finally, by the identification $(\partial K \times \partial T)_+ \simeq A_K$ given by (6), one has $\langle \tau_K(\gamma(t)), p(t) \rangle = -s(t)$, and hence

$$h_T(\dot{C}(t)) = -s(t) + \dot{\lambda}(t).$$

Integrating, we get the desired equality

$$\text{Per}_{h_T}(C) = \int_{S^1} h_T(\dot{C}(t)) dt = \int_{S^1} (-s(t) + \dot{\lambda}(t)) dt = - \int_{\Gamma} s dt,$$

and the proof of the proposition is thus complete. \square

We next recall the definition of the “minimal action” β associated with an invariant circle of a monotone twist map ϕ with generating function h (see e.g., Chapter 1 in [27]):

$$\beta := \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{n=-N}^{N-1} h(r_n, r_{n+1}), \quad (10)$$

where $(r_k, s_k) = \tilde{\phi}^k(r_0, s_0) \in \mathbb{R} \times (-1, 1)$ is the trajectory of an arbitrary point (r_0, s_0) lying on the lift of Γ to $\mathbb{R} \times (-1, 1)$. The minimal action of a convex caustic is defined to be the minimal action of the corresponding invariant circle.

The next proposition relates the minimal action, Lazutkin parameter, perimeter, and rotation number of a convex caustic (see [27] for the Euclidean case, cf. [1]).

Proposition 2.18. *For a symmetric billiard configuration (K, T) , let $C \in \mathfrak{C}(K, T)$ with rotation number ω , minimal action β , and Lazutkin parameter L . Then, one has*

$$L = -\beta - \omega \cdot \text{Per}_{h_T}(C). \quad (11)$$

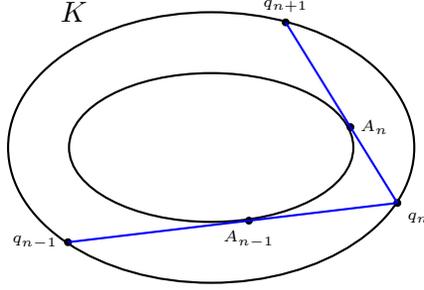


Figure 6

Proof of Proposition 2.18. Denote by Γ the invariant circle associated with the convex T -caustic C . Let (r_n, s_n) , $n \in \mathbb{Z}$, be any $\tilde{\phi}_K$ -trajectory lying on (the lift of) Γ , and let $q_n = \gamma_K(r_n) \in \partial K$ be the corresponding T -billiard trajectory in K . Let $A_n \in C$ be the tangency point between q_n and q_{n+1} (see Figure 6). By the definition of the Lazuktin parameter,

$$\begin{aligned} L &= \text{Per}_{h_T}(\text{Conv}(q_n, C)) - \text{Per}_{h_T}(C) \\ &= h_T(A_n - q_n) + h_T(q_n - A_{n-1}) - \text{Length}_{h_T}(\text{arc}(A_{n-1}A_n)). \end{aligned}$$

Summing over $-N \leq n \leq N-1$ gives

$$2NL = \sum_{n=-N}^{N-1} h_T(q_{n+1} - q_n) + O(1) - \sum_{n=-N}^{N-1} \text{Length}_{h_T}(\text{arc}(A_n A_{n+1})).$$

Thus, using the definition of the generating function (7), and (10), one has

$$\begin{aligned} L &= \frac{1}{2N} \sum_{n=-N}^{N-1} h_T(q_{n+1} - q_n) + O\left(\frac{1}{N}\right) - \frac{1}{2N} \sum_{n=-N}^{N-1} \text{Length}_{h_T}(\text{arc}(A_n A_{n+1})) \\ &\xrightarrow{N \rightarrow \infty} -\beta - \omega \cdot \text{Per}_{h_T}(C). \end{aligned}$$

This completes the proof. \square

Finally we are in a position to prove that dual caustics have equal parameters.

Proof of Proposition 2.16. We first note that the dual caustics $C \subset K$ and $C' \subset T$ have equal rotation numbers. Indeed, this holds as the map $\Psi|_{\Gamma} : \Gamma \rightarrow \Gamma'$ induces an orientation preserving diffeomorphism conjugating the circle homeomorphisms $\phi_K|_{\Gamma}$ and $\phi_T|_{\Gamma'}$, which therefore have the same rotation number (see Chapter 11 in [18]).

Next, we prove that C and C' have equal minimal actions. To this end, we must show that if $(q_n)_{n \in \mathbb{Z}} \subset \partial K$ is a T -billiard trajectory in K which is tangent to C , and $(p_n)_{n \in \mathbb{Z}} \subset \partial T$ is its dual trajectory (which is by definition tangent to C'), then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=-N}^{N-1} h_T(q_n - q_{n+1}) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=-N}^{N-1} h_K(p_n - p_{n+1}). \quad (12)$$

Indeed, if one (and hence both) of the trajectories is periodic, this holds since dual periodic trajectories have equal lengths (see Remark 2.2). The general case now follows, as given two non-periodic dual trajectories, for any fixed N one may close the characteristic curve corresponding to the trajectory $\{(q_n, p_n)\}_{n=-N}^{N-1}$ in $\partial(K \times T)$ with a short segment, obtaining

$$\sum_{n=-N}^{N-1} h_T(q_n - q_{n+1}) = \sum_{n=-N}^{N-1} h_K(p_n - p_{n+1}) + O(1),$$

which proves (12).

From this point we assume that C and C' are C^1 -smooth and strictly convex. We prove next that C and C' have equal perimeters. We write, by a slight abuse of notation, $\Psi: A_K \rightarrow A_T$ for the map induced by Ψ under the identifications $(\partial K \times \partial T)_+ \simeq A_K$ and $(\partial K \times \partial T)_- \simeq A_T$. Since Ψ preserves the standard symplectic form $dq \wedge dp$, and the latter is identified with the form $ds \wedge dt$ on the two cylinders, we deduce that this map is symplectic, i.e., $\Psi^*(ds \wedge dt) = ds \wedge dt$. Note additionally that Ψ extends continuously to a map between the closed cylinders $\Psi: \bar{A}_K \rightarrow \bar{A}_T$, where for example $\bar{A}_K = S^1 \times [-1, 1]$, and this extension maps the boundary circle $\{s = -1\} \subset \bar{A}_K$ to $\{s = -1\} \subset \bar{A}_T$. Now, let Γ and Γ' be invariant circles associated with C and C' , respectively. As these invariant circles are graphs, we may write

$$\Gamma = \{(t, s) \in A_K : s = f(t)\}, \quad \Gamma' = \{(t, s) \in A_T : s = g(t)\}.$$

Consider the domains

$$D = \{(t, s) \in A_K : -1 < s < f(t)\}, \quad D' = \{(t, s) \in A_T : -1 < s < g(t)\}.$$

Since C and C' are dual caustics, $\Psi(\Gamma) = \Gamma'$ by Corollary 2.11. Since, as noted above, Ψ maps the boundary circle $\{s = -1\}$ of \bar{A}_K to that of \bar{A}_T , it follows that $\Psi(D) = D'$. Therefore,

$$\int_D ds \wedge dt = \int_D \Psi^*(ds \wedge dt) = \int_{D'} ds \wedge dt. \quad (13)$$

Using Proposition 2.17 and Stokes' Theorem, one obtains

$$\int_D ds \wedge dt = \int_\Gamma s dt - \int_{\{s=-1\}} s dt = \text{Per}_{h_T}(C) + \int_{S^1} dt = \text{Per}_{h_T}(C) + \text{Per}_{h_T}(K). \quad (14)$$

By a similar computation,

$$\int_{D'} ds \wedge dt = \text{Per}_{h_K}(C') + \text{Per}_{h_K}(T). \quad (15)$$

Finally, since $\text{Per}_{h_T}(K) = \text{Per}_{h_K}(T)$ (see formula (3)), relations (13), (14) and (15) imply that $\text{Per}_{h_T}(C) = \text{Per}_{h_K}(C')$, as required.

To conclude, we observe that C and C' have equal Lazutkin parameters. This follows immediately from (11), and the equality of the other parameters appearing there, which has already been established above. \square

Remark 2.19. Proposition 2.16 is a special case of the fact that the monotone twist maps ϕ_K and ϕ_T have equal marked length spectrum functions (see e.g., [27] for more details).

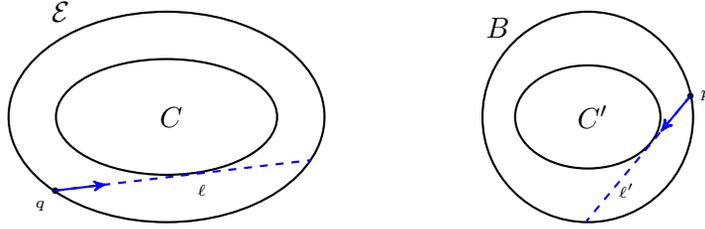


Figure 7. A pair of dual caustics for the Euclidean billiard in an ellipse.

3 Existence of Dual Caustics

In this section we prove Theorem 1.4. We shall start with the special case where the body K is an ellipse. Although this example is quite elementary, it is pivotal in our understanding of Euclidean caustics, since, roughly speaking, every Euclidean string construction is locally an ellipse. This example will play a role both in our proof of the theorem under regularity assumptions on the caustic (Section 3.2), and in the proof provided in the Appendix which pertains to the case of polygonal caustics. The case of a general (Euclidean) caustic is then settled by an approximation argument (Section 3.3).

3.1 The Ellipse Case

It is well known (see e.g., [28]) that confocal ellipses contained in an ellipse \mathcal{E} are convex caustics for the Euclidean billiard in \mathcal{E} . We show, by a straightforward computation, that a Euclidean caustic in an ellipse has a dual caustic in the unit disk B , which is itself an ellipse.

Proposition 3.1. *Let $\mathcal{E} \subset \mathbb{R}_q^2$ be a Euclidean ellipse, given by $A\mathcal{E} = B$ for some positive definite symmetric $A \in \text{GL}(2)$, where B is the Euclidean ball. Let $C \subset \mathcal{E}$ be a confocal ellipse. Then $C' := AC$ is its dual convex \mathcal{E} -caustic in $B \subset \mathbb{R}_p^2$ (see Figure 7).*

While it is relatively easy to check that C' is an \mathcal{E} -caustic in B , using that a linear transformation maps lines to lines and changes length in a predictable way (in fact, for a convex set D , the $h_{\mathcal{E}}$ -perimeter of AD is the Euclidean perimeter of D), Proposition 3.1 asserts something slightly stronger – not only is it a caustic, but it is the caustic *dual to* C . The proof is elementary, and amounts to the following computation.

Proof of Proposition 3.1. Assume without loss of generality that the matrix A is diagonal, i.e.,

$$A = \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix}, \quad \text{with } 0 < b \leq a.$$

In this case

$$\mathcal{E} = \left\{ \frac{q_1^2}{a^2} + \frac{q_2^2}{b^2} \leq 1 \right\}.$$

A confocal ellipse $C \subset \mathcal{E}$ is then of the form

$$C = \left\{ \frac{q_1^2}{a^2 - \lambda} + \frac{q_2^2}{b^2 - \lambda} \leq 1 \right\}, \text{ where } 0 < \lambda < b^2.$$

Then we must show that the dual \mathcal{E} -caustic in B is

$$C' = AC = \left\{ \frac{a^2 p_1^2}{a^2 - \lambda} + \frac{b^2 p_2^2}{b^2 - \lambda} \leq 1 \right\}.$$

Indeed, let $q = (q_1, q_2) \in \partial\mathcal{E}$, and consider a line $\ell = q + \mathbb{R} \cdot p \in \mathcal{L}(\mathcal{E})$ emanating from q with direction $p = (p_1, p_2) \in \partial B$. The line ℓ is mapped under the (\mathcal{E}, B) -duality map α (Definition 2.12) to the line $\ell' = p + \mathbb{R} \cdot n_{\mathcal{E}}(q) \in \mathcal{L}(B)$ (see Figure 7). From Corollary 2.15 it follows that in order to prove our claim, it is enough to check that ℓ is tangent to C if and only if ℓ' is tangent to C' . The intersection $\ell \cap \partial C$ corresponds to solutions $t \in \mathbb{R}$ of the equation

$$\frac{(q_1 + tp_1)^2}{a^2 - \lambda} + \frac{(q_2 + tp_2)^2}{b^2 - \lambda} = 1. \quad (16)$$

Expanding, and using the fact that $q \in \partial\mathcal{E}$, we rewrite (16) as

$$\left(\frac{p_1^2}{a^2 - \lambda} + \frac{p_2^2}{b^2 - \lambda} \right) t^2 + 2 \left(\frac{q_1 p_1}{a^2 - \lambda} + \frac{q_2 p_2}{b^2 - \lambda} \right) t + \lambda \left(\frac{q_1^2}{a^2(a^2 - \lambda)} + \frac{q_2^2}{b^2(b^2 - \lambda)} \right) = 0. \quad (17)$$

Next, note that the outer normal $n_{\mathcal{E}}(q)$ is given, up to rescaling, by the vector $n = \left(\frac{q_1}{a^2}, \frac{q_2}{b^2} \right)$, and so the intersection $\ell' \cap \partial C'$ corresponds to solutions $t \in \mathbb{R}$ of

$$\frac{a^2(p_1 + tq_1/a^2)^2}{a^2 - \lambda} + \frac{b^2(p_2 + tq_2/b^2)^2}{b^2 - \lambda} = 1. \quad (18)$$

Using the fact that $p \in \partial B$, we rewrite (18) as

$$\left(\frac{q_1^2}{a^2(a^2 - \lambda)} + \frac{q_2^2}{b^2(b^2 - \lambda)} \right) t^2 + 2 \left(\frac{q_1 p_1}{a^2 - \lambda} + \frac{q_2 p_2}{b^2 - \lambda} \right) t + \lambda \left(\frac{p_1^2}{a^2 - \lambda} + \frac{p_2^2}{b^2 - \lambda} \right) = 0. \quad (19)$$

Note that the equations (17) and (19) have the same discriminant

$$\Delta = 4 \left(\frac{q_1 p_1}{a^2 - \lambda} + \frac{q_2 p_2}{b^2 - \lambda} \right)^2 - 4\lambda \left(\frac{p_2^2}{a^2 - \lambda} + \frac{p_1^2}{b^2 - \lambda} \right) \left(\frac{q_1^2}{a^2(a^2 - \lambda)} + \frac{q_2^2}{b^2(b^2 - \lambda)} \right).$$

In particular, the condition $\Delta = 0$ is the same for both equations. We deduce that equation (17) has a unique solution if and only if the same holds for equation (19), or in other words, that ℓ is tangent to C if and only if ℓ' is tangent to AC , proving the claim. \square

By a similar computation (or taking the limit $\lambda \nearrow b^2$) we get a dual caustic for the segment between the foci of \mathcal{E} .

Proposition 3.2. *Let $\mathcal{E} \subset \mathbb{R}_q^2$ be an ellipse which is a Euclidean string construction over the segment $C := [(-x, 0), (x, 0)]$, for $x > 0$, with string length $L + 2x$. Then, the segment $C' := [(-\frac{x}{2L}, 0), (\frac{x}{2L}, 0)]$ is the dual convex \mathcal{E} -caustic to C in $B \subset \mathbb{R}_p^2$. Moreover, if $\ell \in \mathcal{L}(\mathcal{E})$ passes through $(x, 0)$, then the (\mathcal{E}, B) -dual line $\alpha_{\mathcal{E}, B}(\ell) \in \mathcal{L}(B)$ passes through $(\frac{x}{2L}, 0)$.*

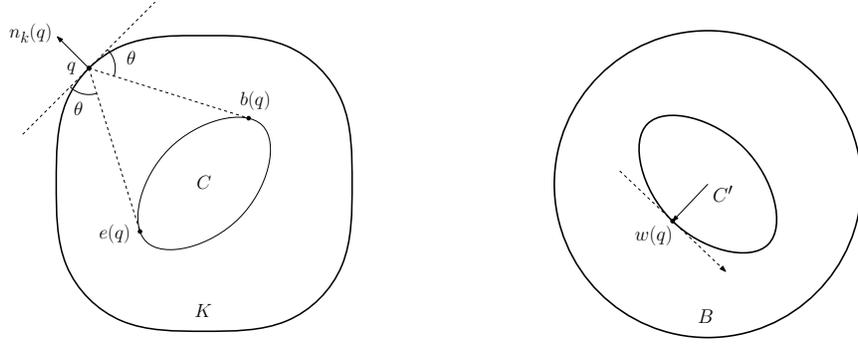


Figure 8. A pair of dual caustics $C \in \mathfrak{C}(K, B)$ and $C' \in \mathfrak{C}(B, K)$.

Using the rotational symmetry of B , we obtain the same result for a rotated ellipse.

Corollary 3.3. *Let $\mathcal{E} \subset \mathbb{R}_q^2$ be an ellipse which is a Euclidean string construction over the segment $C := [e, b]$ with string length $L + |e - b|$, for two distinct points $b, e \in \mathbb{R}_q^2$. Then $C' := [\frac{e-b}{2L}, \frac{b-e}{2L}]$ is its dual convex \mathcal{E} -caustic in $B \subset \mathbb{R}_p^2$. Moreover, if $\ell \in \mathcal{L}(\mathcal{E})$ passes through e (b), then the (\mathcal{E}, B) -dual line $\alpha_{\mathcal{E}, B}(\ell) \in \mathcal{L}(B)$ passes through $\frac{e-b}{2L}$ ($\frac{b-e}{2L}$).*

This result will be crucial in the proof of Theorem 1.4 below, and sheds some light on the formula for the dual caustic given in Remark 1.5.

3.2 The case of smooth caustics

Here we prove a special case of Theorem 1.4 when the caustic C is sufficiently regular.

Let K and C be as in Theorem 1.4, and assume moreover that C is strictly convex. For $q \in \partial K$, denote by $e(q)$ and $b(q)$ the positive and the negative tangency points to C from q (see Figure 8), and set $L(q) = |q - e(q)| + |q - b(q)|$.

Proposition 3.4. *Let $K \subset \mathbb{R}_q^2$ be a C^1 -smooth, centrally symmetric and strictly convex body. If $C \subset K$ is a Euclidean caustic which is C^2 -smooth and has nowhere vanishing curvature, then it admits a dual K -caustic $C' \subset B$. Moreover, C' is C^1 -smooth, strictly convex, and its boundary can be parametrized by*

$$w(q) = \frac{e(q) - b(q)}{L(q)}, \quad q \in \partial K.$$

For the proof of Proposition 3.4 we will require some straightforward computations. In what follows we denote by τ_C the unit counter-clockwise tangent to C and by k_C the curvature of C . Moreover, we denote by θ the angle of incidence and reflection at q (see Figure 8), namely,

$$\cos \theta = \langle \tau_K(q), \frac{e - q}{|e - q|} \rangle = \langle \tau_K(q), \frac{q - b}{|q - b|} \rangle.$$

Lemma 3.5. *With the assumptions of Proposition 3.4, the functions $e(q)$, $b(q)$, and $L(q)$ are C^1 -smooth, with derivatives (with respect to the counter-clockwise arc-length parameter along the boundary ∂K)*

$$\begin{aligned} e'(q) &= \frac{\sin \theta}{k_C(e)|e - q|} \tau_C(e), & b'(q) &= \frac{\sin \theta}{k_C(b)|b - q|} \tau_C(b), \\ L'(q) &= \sin \theta \cdot \left[\frac{1}{k_C(e)|q - e|} - \frac{1}{k_C(b)|q - b|} \right]. \end{aligned}$$

Proof of Lemma 3.5. Note that the function $e(q)$ is determined implicitly by q via the following equation

$$\langle q - e(q), n_C(e(q)) \rangle = 0.$$

In fact, this equation has two solutions, one of which is $e(q)$ and the other $b(q)$. Let $\gamma : [0, A] \rightarrow \partial C$ be a counter-clockwise parametrization by (Euclidean) arc length. We rewrite the above equation as

$$F(q, t) = \langle q - \gamma(t), n_C(\gamma(t)) \rangle = 0.$$

Note that $F : \partial K \times [0, A] \rightarrow \mathbb{R}$ is C^1 -smooth. To see that e depends smoothly on q , we must check that $\frac{\partial F}{\partial t} \neq 0$. Indeed,

$$\frac{\partial F}{\partial t}(q, t) = \underbrace{-\langle \gamma'(t), n_C(\gamma(t)) \rangle}_{=0} + \langle q - \gamma(t), k_C(\gamma(t)) \dot{\gamma}(t) \rangle = k_C(\gamma(t)) \cdot \langle q - \gamma(t), \dot{\gamma}(t) \rangle.$$

Note that by the definition of e , the vector $q - \gamma(t)$ is negatively tangent to C at $\gamma(t)$, so that $q - \gamma(t) = -|q - \gamma(t)| \dot{\gamma}(t)$. Using the assumption $k_C \neq 0$, we conclude that for this solution (q, t) of the equation $F(q, t) = 0$ one has

$$\frac{\partial F}{\partial t}(q, t) = -|q - \gamma(t)| \cdot k_C(\gamma(t)) \cdot \langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle = -|q - \gamma(t)| \cdot k_C(\gamma(t)) \neq 0.$$

This shows that the map $q \mapsto t(q)$ for the above chosen solution of $F(q, t(q)) = 0$ is C^1 -smooth. We next compute its derivative with respect to q . Clearly,

$$\frac{\partial F}{\partial q}(q, t) = \langle \tau_K(q), n_C(\gamma(t)) \rangle = -\langle n_K(q), \tau_C(\gamma(t)) \rangle = -\cos(\theta + \pi/2) = \sin \theta.$$

Then, the Implicit Function Theorem gives

$$\frac{dt}{dq} = -\frac{\partial F/\partial q}{\partial F/\partial t} = \frac{\sin \theta}{k_C(\gamma(t))|\gamma(t) - q|}.$$

Letting $e(q) = \gamma(t(q))$, we conclude that e depends C^1 -smoothly on q , and

$$e'(q) = \gamma'(t(q)) \cdot \frac{dt}{dq} = \frac{\sin \theta}{k_C(e(q))|e(q) - q|} \tau_C(e(q)).$$

In computing $b'(q)$, the only change is that the solution t of $F(q, t) = 0$ is now chosen so that $q - \gamma(t)$ is positively tangent to C at $\gamma(t)$, which yields identical computations except for the following signs:

$$q - \gamma(t) = |q - \gamma(t)| \dot{\gamma}(t) \quad \text{and} \quad \langle n_K(q), \tau_C(\gamma(t)) \rangle = \cos(\pi/2 - \theta) = \sin \theta.$$

Finally, since $e(q)$ and $b(q)$ are C^1 -smooth, so is $L(q)$. Elementary differentiation gives

$$L'(q) = \frac{\langle e(q) - q, e'(q) \rangle}{|e(q) - q|} + \frac{\langle b(q) - q, b'(q) \rangle}{|b(q) - q|} = \langle e'(q), \tau_C(e(q)) \rangle - \langle b'(q), \tau_C(b(q)) \rangle.$$

Substituting the formulas for $e'(q)$ and $b'(q)$ gives the required result. \square

Proof of Proposition 3.4. By the previous lemma, the map

$$\partial K \ni q \mapsto w(q) = \frac{e(q) - b(q)}{L(q)}$$

is a closed, C^1 -smooth parametrized curve which is clearly contained in B . Our goal is to prove that it is a simple closed curve bounding a strictly convex set, which we denote by C' , and that $C' \subset B$ is a K -caustic which is dual to C . To this end, we first prove that $w'(q)$ is negatively proportional to $n_K(q)$ for all $q \in \partial K$. We first fix some notations. We abbreviate $e = e(q)$, $b = b(q)$, and denote

$$B = |b - q|, \quad v = \frac{b - q}{B}, \quad E = |e - q|, \quad u = \frac{e - q}{E}.$$

Next, by the definition of e and b , one has that $\tau_C(e) = u$ and $\tau_C(b) = -v$. Note that by the (Euclidean) billiard reflection law, $n_K(q)$ is negatively proportional to the vector $u + v$ (see Figure 8). A direct computation using Lemma 3.5 gives

$$w'(q) = \frac{\sin \theta}{L(q)^2} \left[\frac{B}{E \cdot k_C(e)} + \frac{E}{B \cdot k_C(b)} \right] (u + v).$$

This proves that $w'(q)$ is proportional to $-n_K(q)$, as asserted. This implies that the angle of $w'(q)$ is strictly increasing, and increases by 2π as q traverses ∂K (counter-clockwise) once. Finally, we use elementary differential geometry to conclude that $w(q)$ is a simple closed curve bounding a strictly convex domain C' . Indeed, simplicity follows since otherwise w would self-intersect at some point p , in which case either the intersection is transversal, which would imply that points of w lie on both sides of one of the tangents to w at p , or the intersection is tangential, in which case w would have a double tangent at p . In both cases, this implies the existence of three parallel tangents to w , a contradiction. For the convexity, note that once the curve is simple, if it were not convex, this would mean that we may construct a line which intersects w at least four times, partitioning it into four arcs, and such that w admits a tangent parallel to this line in each of the four regions, again contradicting the fact that the angle increases by 2π during the journey along w .

Next let us prove that C' is the dual caustic to C . Indeed, fix $q \in \partial K$ and let \mathcal{E} be the ellipse formed by the Euclidean string construction over the segment $[e(q), b(q)]$ with string length $L(q) + |e(q) - b(q)|$. Note that \mathcal{E} contains q on its boundary, and moreover $n_{\mathcal{E}}(q) = n_K(q)$ (since the broken line \overline{bqe} obeys the Euclidean billiard reflection law in both). In particular, the duality transforms $\alpha_{K,B}$ and $\alpha_{\mathcal{E},B}$ agree on oriented lines emanating from q . By Corollary 3.3, the dual of the oriented line ℓ emanating from q and tangent to C at $e(q)$, $\alpha_{\mathcal{E},B}(\ell)$, passes through $w(q)$, and has direction $n_{\mathcal{E}}(q) = n_K(q)$. Since $w'(q)$ is

negatively proportional to $n_K(q)$, it follows that the line $\alpha_{K,B}(\ell) = \alpha_{\varepsilon,B}(\ell)$ is tangent to C' . In other words, denoting by Γ the invariant circle determined by the caustic $C \subset K$, we have shown that C' is tangent to every line in $\alpha_{K,B}(\Gamma)$. By Corollary 2.15, this means that C' is the dual convex caustic to C . This completes the proof. \square

3.3 Proof of Theorem 1.4

Here we complete the proof of Theorem 1.4, by using an approximation argument to reduce it to Proposition 3.4. To this end we shall need the following lemma.

Lemma 3.6. *Suppose that $K_n \subset \mathbb{R}_q^2$ and $T_n \subset \mathbb{R}_p^2$ are a sequence of symmetric billiard configurations, and let C_n be a sequence of convex T_n -caustics in K_n . Assume further that the following Hausdorff limits exist:*

$$K_n \rightarrow K, T_n \rightarrow T, C_n \rightarrow C,$$

where (K, T) is a symmetric billiard configuration. Then, C is a (convex) T -caustics in K . Moreover, if in addition for each $n \in \mathbb{N}$ there exists a convex K_n -caustic C'_n dual to C_n , and $C'_n \rightarrow C'$ in the Hausdorff metric, then C and C' are dual caustics.

Proof of Lemma 3.6. The fact that $C \subseteq K$ follows from the properties of Hausdorff convergence. To show that C is a T -caustic in K , note that any positive tangent line ℓ to C is a limit of positive tangent lines ℓ_n to C_n . Indeed, let $x \in \partial K$. From the fact that $K_n \rightarrow K$ it follows that there is a sequence $x_n \in \partial K_n$ such that $x_n \rightarrow x$. For each x_n let ℓ_n be the positive tangent to C_n emanating from x_n . By passing to a subsequence, we may assume without loss of generality that ℓ_n converges to some oriented line $\tilde{\ell}$ emanating from the point x . It is easily seen that the line $\tilde{\ell}$ is tangent to C . This follows from the fact that the tangency of ℓ_n to C_n can be described by the equality $\max_{C_n} \phi_n = \phi_n(x_n)$, where ϕ_n is a linear functional (say, of norm one) which admits ℓ_n as a level set. Since the line emanating from x and positively tangent to C is unique, one has $\tilde{\ell} = \ell$ i.e., $\ell_n \rightarrow \ell$. To show that C is indeed a T -caustic, consider for each n the next reflection point $x'_n \in \partial K_n$ of the line ℓ_n , and the line ℓ'_n emanating from x'_n and positively tangent to C_n . In other words, $\Psi_{K_n, T_n}^2(\ell_n) = \ell'_n$. Repeating the argument above, we obtain that $\ell'_n \rightarrow \ell'$ which is positively tangent to C . From Corollary 2.14 it follows that $\Psi_{K, T}^2(\ell) = \ell'$, which shows that C is indeed a T -caustic in K . For the second part of the lemma, let C'_n be a convex K_n -caustic in T_n such that $C'_n \rightarrow C'$. As before, C' is a K -caustic in T . Moreover, since $\Psi_{K_n, T_n}(\ell_n)$ is tangent to C'_n , passing to the limit one sees that $\Psi_{K, T}(\ell)$ is tangent to C' , and so by Corollary 2.11, C and C' are dual caustics. \square

Proof of Theorem 1.4. Let K be a centrally symmetric, smooth, and strictly convex body, and let C be a Euclidean caustic in K . By Lemma 2.8 the body K is a (Euclidean) string construction over C , and we denote its string length by L . Let C_n , $n \in \mathbb{N}$, be a sequence of C^2 -smooth convex bodies with nowhere vanishing curvature which converges to C in the Hausdorff metric. For each n , let K_n be the convex body formed by a Euclidean

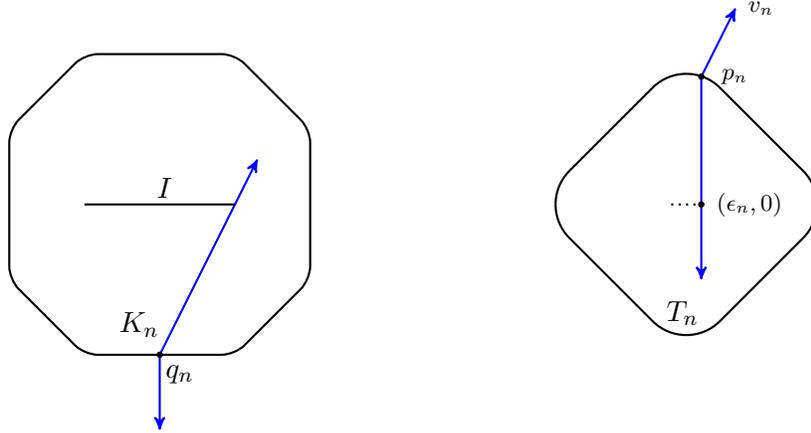


Figure 9. A convex caustic which has no dual convex caustic.

string construction over C_n with string length L . Then, denoting $f_n(x) = \text{Per}(\text{Conv}(x, C_n))$ and $f(x) = \text{Per}(\text{Conv}(x, C))$, one has $K_n = \{f_n \leq L\}$ and $K = \{f \leq L\}$. Since $C_n \rightarrow C$, one easily verifies that (f_n) converges to f uniformly, which implies that K_n converge to K in the Hausdorff metric. By Proposition 3.4, each C_n admits a dual K_n -caustic $C'_n \subset B$. By passing to a subsequence, we may assume that C'_n converge in the Hausdorff metric to a convex body C' . Now by Lemma 3.6, C' is a K -caustic in B , dual to C . By Proposition 2.16, the caustics C and C' have the same rotation numbers, and that the dual caustics C_n and C'_n have equal perimeters and string lengths. Finally, by continuity, C and C' have equal perimeters and equal string lengths. This completes the proof. \square

4 Non-existence of a dual caustic

In this section we prove Theorem 1.6 and show that the caustic-to-caustic duality established in Theorem 1.4 fails for arbitrary Minkowski billiards.

Proof of Theorem 1.6. We start with a construction of a “degenerate” convex caustic which has no convex dual caustic. Let $T \subset \mathbb{R}_p^2$ be the unit ball of the ℓ_1 -norm, i.e., $T = \{(p_1, p_2) : |p_1| + |p_2| \leq 1\}$. Let T_n be a family of smooth, unconditional, and strictly convex bodies in \mathbb{R}_p^2 which converges to T in the Hausdorff metric. Assume further that for each n , the boundary of T_n contains the point $(1, 0)$, and that the (Euclidean) outer normal n_{T_n} to T_n on the open first quadrant converges locally uniformly to the vector $(1/\sqrt{2}, 1/\sqrt{2})$ (and hence by unconditionality $n_{T_n} \rightarrow (\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$ locally uniformly on the other open quadrants). Fix the interval $I = [-1, 1] \subset \mathbb{R}_q^2$, and a string length $L = 6$, and let $K_n \subset \mathbb{R}_q^2$ be the result of the h_{T_n} -string construction on I with string length L (see Figure 9). By Lemma 2.8, $I = [-1, 1]$ is a convex T_n -caustic in K_n for all n . We claim that for large enough n , the caustic I admits no dual convex caustic in T_n .

Indeed, assume towards a contradiction that for infinitely many n there exists a convex

K_n -caustic C_n in T_n , dual to I . We first note that the caustic $I \subset K_n$ admits a 2-periodic tangent T_n -billiard trajectory along the q_1 -axis, with bouncing points $(\pm 2, 0)$. Since the vector $n_{K_n}(\pm 2, 0) = (\pm 1, 0)$, the dual caustic C_n also admits a 2-periodic tangent K_n -billiard trajectory along the p_1 -axis, and hence C_n must be “flat”, that is $C_n \subset \{p_2 = 0\}$. From symmetry, we conclude that $C_n = [-\epsilon_n, \epsilon_n] \times \{0\} \subset \mathbb{R}_p^2$, for some $\epsilon_n > 0$. Next, let $q_n = (0, -y_n) \in \partial K_n$ with $y_n > 0$. Note that, by symmetry, $n_{K_n}(q_n)$ is in direction $(0, -1)$. Let ℓ_n be the positive tangent from q_n to I , whose direction we denote by v_n (normalized so that $v_n \in \partial T_n$). Then, by Corollary 2.15, the line $\alpha_{K_n, T_n}(\ell_n)$ is (negatively) tangent to C_n . By definition, $\alpha_{K_n, T_n}(\ell_n) = (p_n, (0, -1))$, where $p_n \in \partial T_n$ is the unique point such that $n_{T_n}(p_n) = v_n$ (see Figure 9). Therefore, the first coordinate of p_n is ϵ_n . Note that the sequence v_n converges to some $v \notin \{(t, \pm t) : t \in \mathbb{R}\}$. Indeed, K_n converges to K , which is the body obtained by the h_T -string construction on the interval I , and so $v_n \rightarrow v$, which is the vector in the direction of the positive tangent to I from $q = (0, -2) \in \partial K$, i.e., v is in direction $(1, 2)$. Since $v_n = n_{T_n}(p_n) \rightarrow v$, by the construction of T_n one has that $p_n = n_{T_n^\circ}(v_n) \rightarrow n_{T^\circ}(v) = (0, 1)$. This shows that $\epsilon_n \rightarrow 0$.

Using Lemma 2.8, as $C_n = [-\epsilon_n, \epsilon_n] \times \{0\}$ is a K_n -caustic in T_n , it is an h_{K_n} -string construction on $[-\epsilon_n, \epsilon_n] \times \{0\}$ of some string length L_n . In fact, since ∂T_n contains the point $(1, 0)$, one has that $L_n = 4(1 + \epsilon_n)$ (as $h_{K_n}(1, 0) = 2$). We claim that from this it follows that the Hausdorff distance $d_H(T_n, (L_n/2) \cdot K_n^\circ) \rightarrow 0$ as $n \rightarrow \infty$. Indeed, by item (i) in Lemma 2.9, it follows that T_n is contained in $(L_n/2) \cdot K_n^\circ$. On the other hand, let us verify that T_n contains the result of an h_{K_n} -string construction of string length $L_n - 8\epsilon_n$ (which is positive as $\epsilon_n < 1$) on the point $\{(0, 0)\}$, which is $(L_n/2 - 4\epsilon_n) \cdot K_n^\circ$. Indeed, let p belong to $(L_n/2 - 4\epsilon_n) \cdot K_n^\circ$, so that $h_{K_n}(p) \leq L_n/2 - 4\epsilon_n$. By the triangle inequality, and as $h_{K_n}(\epsilon_n, 0) = 2\epsilon_n$, one has

$$h_{K_n}(p - (\epsilon_n, 0)) \leq h_{K_n}(p) + 2\epsilon_n \leq L_n/2 - 2\epsilon_n.$$

Similarly, $h_{K_n}(p + (\epsilon_n, 0)) \leq L_n/2 - 2\epsilon_n$. It follows that the h_{K_n} -perimeter of $\text{Conv}(p, C_n)$ is at most L_n , so $p \in T_n$. Since $L_n = 4(1 + \epsilon_n)$, we verified that

$$(2 - 2\epsilon_n) K_n^\circ \subset T_n \subset (2 + 2\epsilon_n) K_n^\circ.$$

Therefore, the Hausdorff distance $d_H(T_n, 2K_n^\circ) \rightarrow 0$ as $n \rightarrow \infty$. By construction, $T_n \rightarrow T$ as $n \rightarrow \infty$. Therefore, $K_n \rightarrow 2T^\circ$, which is a square of side length 4. On the other hand, by the continuity of the string construction, K_n converge to the h_T -string construction of length $L = 6$ on I , which is clearly not a square (it is an octagon, see Figure 9). We have thus reached a contradiction, meaning that for large enough n , the T_n -caustic I in K_n has no dual convex caustic.

To produce a counterexample with a smooth caustic we use a standard approximation argument. Let (K, T) be a symmetric billiard configuration admitting an interval $I \subset K$ as a T -caustic with string length L which has no dual convex K -caustic (as constructed above). Let \mathcal{E}_j be a sequence of ellipses such that $\mathcal{E}_j \rightarrow I$, and let $K_j \subset \mathbb{R}_q^2$ be the result of the h_T -string construction on \mathcal{E}_j with string length L . We claim that for j large enough, T contains no convex K_j -caustic dual to \mathcal{E}_j . Indeed, if a dual caustic exists for infinitely

many j 's, then from Lemma 3.6 it follows that in the limit (after passing to a subsequence if necessary) I admits a convex K -dual caustic in T which is a contradiction. \square

5 Appendix: dual caustics to polygonal caustics

To complete the picture we present a different proof for Theorem 1.4, using an approximation of a general B -caustic in K by polygons. The approximation argument is similar to the one in the proof of Theorem 1.4, using Lemma 3.6. The main ingredient is the following theorem for polygonal caustics.

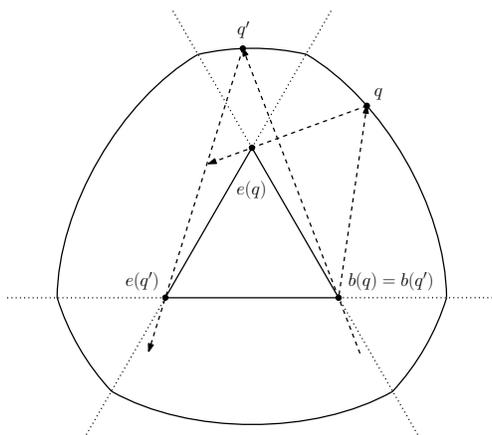


Figure 10. String construction over a triangle.

Theorem 5.1. *Let C be a convex polygon which is a caustic for the Euclidean billiard in a C^1 -smooth, centrally symmetric strictly convex body $K \subset \mathbb{R}_q^2$. Then C has a convex dual polygonal caustic $C' \subset B^2 \subset \mathbb{R}_p^2$.*

Proof of Theorem 5.1. First note that since K is a Euclidean string construction over C , then by Lemma 2.7 it follows that K is necessarily C^1 -smooth and strictly convex. Next, for $q \in \partial K$, we denote by $e(q)$ and $b(q)$ the positive and negative tangency points to C from q , and let $L(q) = |q - e(q)| + |q - b(q)|$. These notation are as in the proof of Proposition 3.4, but the main difference is that in the polygonal case the functions $e(q)$ and $b(q)$ are no longer continuous, and are piecewise constant (and initially multi-valued at a finite number of points). In fact, the boundary ∂K is a finite union of arcs, on (the interior of) each of which the functions $e(q)$ and $b(q)$ are constant, and those arcs are therefore arcs of ellipses with foci $e(q)$ and $b(q)$ and string length given by $L(q) + |e(q) - b(q)|$ (see Figure 10)[†].

Let us explain this in more detail. Note that, as q traverses ∂K counter-clockwise, the changes in $b(q)$ and $e(q)$ occur only when q crosses the lines passing through the edges of C , and at such an intersection either one or both of $e(q)$ and $b(q)$ change, by moving from one vertex of C to a successive vertex (and for a generic string length they do not change

[†]Note that a string construction over a non-centrally-symmetric polygon can never be centrally symmetric. However, the figure suffices to illustrate the parameters specified.

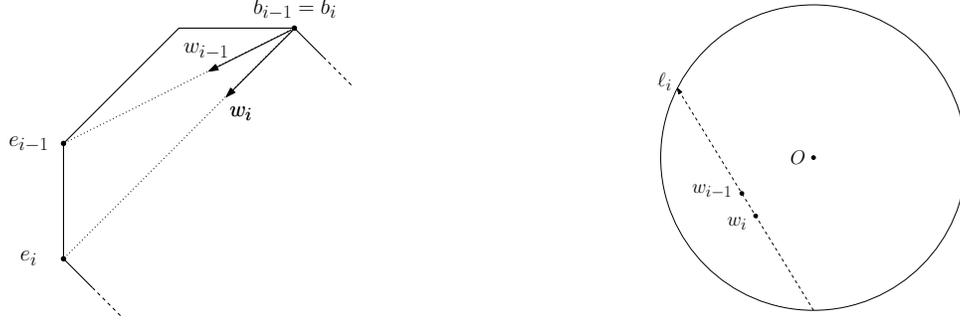


Figure 11

simultaneously). Indeed, for any point $z \in \mathbb{R}^n$ one may consider the edges of C which are “illuminated” by z and those which are “dark”. Here an edge is called “illuminated by z ” if the interval $[z, q] \cap C = \{q\}$ for any q on that edge. It is not hard to check that for any fixed $z \in \mathbb{R}^2 \setminus C$, the edges of C are separated into two disjoint sets (illuminated and dark). Moreover, with respect to the cyclic order on the edges, the illuminated edges form an order interval, as do the dark edges. For $q \in \partial K$, the point $b(q)$ is the initial vertex of the illuminated polygonal line, and $e(q)$ is its terminal vertex, and we use this to extend the definitions of $e(q)$ and $b(q)$ to those points where they were previously multi-valued. As q crosses the line on which a certain edge lies, this edge changes from illuminated to dark, or vice versa. In the former case, $b(q)$ changes from the initial vertex of this edge to its terminal vertex, and in the latter case, $e(q)$ changes from the initial vertex of this edge to its terminal vertex (again, see Figure 10). This shows that, as asserted, the boundary of K is a finite union $\partial K = \bigcup_{i=1}^M \mathcal{E}_i$, disjoint except for common boundary points, where for each i the functions $e(q)$ and $b(q)$ are constant on the interior of \mathcal{E}_i , and \mathcal{E}_i is therefore an arc of an ellipse with foci $b(q)$ and $e(q)$. We index \mathcal{E}_i according to the counter-clockwise order on ∂K , and denote by q_i the joint boundary point of \mathcal{E}_{i-1} and \mathcal{E}_i , and by e_i , b_i and L_i the constant values of $e(q)$, $b(q)$, and $L(q)$ on the interior of \mathcal{E}_i .

Denote, for $1 \leq i \leq M$, $w_i = \frac{e_i - b_i}{L_i} \in B$, and consider the closed polygonal line P with vertices $(w_i)_{i=1}^M$, in that order. We will prove that P is a simple closed polygonal line, bounding a convex polygon C' , and that C' is in fact a convex K -caustic in B dual to C . Indeed, consider the invariant circle Γ corresponding to C . The partition $\partial K = \bigcup_i \mathcal{E}_i$ induces a partition $\Gamma = \bigcup_i \Gamma_i$, where for each i , Γ_i corresponds to those oriented lines of Γ emanating from points of \mathcal{E}_i . We have observed that \mathcal{E}_i is an arc of an ellipse with foci b_i and e_i and string length $L_i + |e_i - b_i|$, and Γ_i contains those lines emanating from points of \mathcal{E}_i and passing through e_i . Therefore, by Corollary 3.3, the image of Γ_i under the (K, B) -duality map α contains lines emanating from some arc of ∂B , which we denote by B_i , and passing through w_i .

For each i we consider the point $p_i = B_i \cap B_{i-1} \in \partial B$. Denote by $\ell_i \in \Gamma$ the oriented line emanating from q_i , and $\ell'_i = \alpha(\ell_i) \in \alpha(\Gamma)$, which emanates from p_i . Note that ℓ'_i is in direction $n_K(q_i)$, and that it belongs to both $\alpha(\Gamma_{i-1})$ and $\alpha(\Gamma_i)$, hence it passes through both w_{i-1} and w_i , so $w_i - w_{i-1}$ is parallel to the outer normal $n_K(q_i)$. We claim that it

is in fact negatively proportional to this normal. Indeed, by Corollary 3.3, the line ℓ'_i is negatively tangent to the segment whose end-point is w_i , and hence the origin O lies to the right of ℓ'_i . Additionally, one easily verifies that (w_{i-1}, w_i) forms a positive basis of \mathbb{R}^2 , since by construction w_{i-1} and w_i either share a common initial point b_i , in which case e_{i-1} and e_i are consecutive vertices of C and b_i lies in C , or they share the end point e_i and b_{i-1} and b_i are consecutive points in C (see Figure 11). It follows easily that $w_{i-1} - w_i$ agrees with the orientation of ℓ'_i , which means that $w_i - w_{i-1}$ is negatively proportional to $n_K(q_i)$ (see Figure 11 again). It follows that the angle of the edges $w_i - w_{i-1}$ is strictly increasing, and increases by 2π , as i runs from 1 to M . Similarly to the smooth case, this implies that the polygonal line P is a simple closed curve, which bounds a convex polygon C' . Indeed, simplicity follows as before, since at a self-intersection there must occur one of the following two: (1) at least one of the (inbound or outbound) edges, when elongated to a line, is such that P lies on both of its sides, or (2) the intersection is a segment along an edge; both cases lead to a contradiction with the fact that the angle of edges increases only by 2π during the journey along P . For the convexity one repeats the argument in the smooth case verbatim.

We are left with showing that C' is the dual caustic to C . Indeed, by construction the dual invariant circle $\alpha(\Gamma)$ contains lines which pass through the vertices of C' . By the convexity of C' , this means that C' is everywhere tangent to $\alpha(\Gamma)$. By Corollary 2.15, this means that C' is a convex caustic dual to C , as required. \square

References

- [1] Amiran, E. Y. *A dynamical approach to symplectic and spectral invariants for billiards*, Comm. Math. Phys. 154 (1993), no. 1, 99–110.
- [2] Artstein-Avidan, S., Ostrover, Y. *Bounds for Minkowski billiard trajectories in convex bodies*, Inter. Math. Res. Not. Vol. 2014, (2012) no. 1, 165–193.
- [3] Balestro, V., Martini, H., Shonoda, E. *Concepts of curvatures in normed planes*, arXiv:1702.01449.
- [4] Berger, M. *Seules les quadriques admettent des caustiques*, Bull. Soc. Math. France 123 (1995), 107–116.
- [5] Bialy, M. *Convex billiards and a theorem by E. Hopf*, Math. Z. 214 (1993), no. 1, 147–154.
- [6] Biberstein, O. A. *Elements de géométrie différentielle minkowskienne*, PhD thesis, Université de Montreal, 1957.
- [7] Birkhoff, G. *Dynamical Systems*, Volume 9, American Mathematical Society Colloquium Publications, 1927.
- [8] Chernov, N., Markarian, R. *Chaotic Billiards*, in: Mathematical Surveys and Monographs, vol. 127, American Mathematical Society, Providence, RI, 2006.

- [9] Douady, R. *Application du théorème des tores invariants. Thèse de troisième cycle*, Univ. Paris 7 (1982).
- [10] Gruber, P. M. *Convex billiards*, 33 Geom. Ded. (1990) 205–226.
- [11] Gruber, P. M. *Only ellipsoids have caustics*, Math. Ann. 303 (1995), 185–194.
- [12] Gutkin, E. *Billiard dynamics: an updated survey with the emphasis on open problems*, Chaos 22 (2012), no. 2, 026116.
- [13] Gutkin, E., Katok, A. *Caustics for inner and outer billiards*, Comm. Math. Phys. 173 (1995), 101–133.
- [14] Gutkin, E., Tabachnikov, S. *Billiards in Finsler and Minkowski geometries*, J. Geom. Phys. 40 (2002), no. 3-4, 277–301.
- [15] Halpern, B. *Strange billiard tables*, Trans. Amer. Math. Soc. 232: (1977) 297–305.
- [16] Hubacher, A. *Instability of the boundary in the billiard ball problem*, Comm. Math. Phys. 108 (1987), no. 3, 483–488.
- [17] Joets, A., Ribotta, R. *A geometrical model for the propagation of rays in an anisotropic inhomogeneous medium*, Optics Commun. Volume 107 (1994), Issues 3-4, 200–204.
- [18] Katok, A., Hasselblatt, B. *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge, 1995.
- [19] Knill, O. *On nonconvex caustics of convex billiards*, Elem. Math. 53 (1998), no. 3, 89–106.
- [20] Kozlov, V., Treshchëv, D. *Billiards*, Amer. Math. Soc., Providence RI, 1991.
- [21] Lazutkin, V. F. *Existence of caustics for the billiard problem in a convex domain*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 37 (1973), 186–216.
- [22] Lazutkin, V.F. *KAM Theory and Semiclassical Approximations to Eigenfunctions*, New York: Springer-Verlag, 1993.
- [23] Mather, J. N. *Glancing billiards*, Ergodic Theory Dyn. Syst. 2 (1982), no. 3-4, 397–403.
- [24] Petty, C. M. *On the geometry of the Minkowski plane*, Riv. Mat. Univ. Parma 6 (1955), 269–292.
- [25] Rockafellar, R. T. *Convex Analysis*, Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997.
- [26] Schneider, R. *Convex bodies: the Brunn-Minkowski Theory*, Second expanded edition. Encyclopedia of Mathematics and its Applications, 151. Cambridge University Press, Cambridge, 2014.

- [27] Siburg, K. F. *The Principle of Least Action in Geometry and Dynamics*, Lecture Notes in Mathematics Vol. 1844, Springer-Verlag, 2004.
- [28] Tabachnikov, S. *Geometry and Billiards*, Student Mathematical Library, 30. American Mathematical Society, Providence, RI; University Park, PA, 2005.
- [29] Tabachnikov, S. *Billiards*, SMF Panoramas et Synthèses, No. 1, 1995.
- [30] Thompson, A.C. *Minkowski Geometry*, Encyclopedia of Mathematics and its Applications, 63. Cambridge University Press, Cambridge, 1996.

Shiri Artstein-Avidan
School of Mathematical Sciences
Tel Aviv University, Tel Aviv 69978, Israel
e-mail: artstein@post.tau.ac.il

Dan Florentin
Department of Mathematical Sciences
Kent State University, Kent, OH, 44242 USA
e-mail: danflorentin@gmail.com

Yaron Ostrover
School of Mathematical Sciences
Tel Aviv University, Tel Aviv 69978, Israel
e-mail: ostrover@post.tau.ac.il

Daniel Rosen
School of Mathematical Sciences
Tel Aviv University, Tel Aviv 69978, Israel
e-mail: danielr6@post.tau.ac.il